



Universidad de Granada

Papel del elemento Line-1 en la patología humana de anemia Fanconi

Santiago Morell Hita

Programa oficial de doctorado en Biomedicina

Editor: Editorial de la Universidad de Granada
Autor: Santiago Morell Hita
D.L.: GR 261-2015
ISBN: 978-84-9083-282-0

El doctorando Santiago Morell Hita y el director de la tesis José Luis García Pérez, garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección del director de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, a 5 de Septiembre de 2014

Director de la Tesis:

Jose Luis Garcia Perez

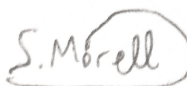
Fdo.:



Doctorando:

Santiago Morell Hita

Fdo.:



INDICE DE CONTENIDOS

INTRODUCCIÓN Y OBJETIVOS

1.- INTRODUCCIÓN	1
1.1.- Genomas y elementos móviles	1
1.1.1.- Complejidad de los genomas	1
1.1.2.- Elementos móviles	3
1.1.3.- Clases de elementos móviles	6
1.1.3.1.- DNA transposones	10
1.1.3.2.- Retrotransposones	14
1.1.3.2.1.- Intrones tipo II	18
1.1.3.2.2.- Retrotransposones LTR	20
1.1.3.2.3.- Retrotransposones No-LTR	24
1.1.4.- Elementos móviles en el genoma humano	26
1.2.- Los elementos LINE-1 (L1) en el genoma humano	30
1.2.1.- Descubrimiento	30
1.2.2.- Estructura	32
1.2.3.- Mecanismo de movilización de LINE-1: TPRT	36
1.2.4.- Consecuencias de la actividad de LINE-1	39
1.2.5.- Regulación de LINE-1	43
1.2.6.- Vía de retrotransposición de LINE-1 independiente de la endonucleasa (ENi) y reparación del DNA	48
1.3.- Anemia de Fanconi (<i>Fanconi Anemia</i> o FA)	53
1.3.1.- Introducción	53
1.3.2.- Sintomatología	55
1.3.3.- Diagnóstico y tratamiento	58
1.3.4.- Etiología y bases moleculares: ruta de la Anemia de Fanconi	62
1.3.5.- Miembros de la ruta de la Anemia De Fanconi	64
1.3.5.1.- FANCA	69
1.3.5.2.- FANCB	69
1.3.5.3.- FANCC	70
1.3.5.4.- FANCD1 (BRCA2)	70
1.3.5.5.- FANCD2	72
1.3.5.6.- FANCE	74
1.3.5.7.- FANCF	75
1.3.5.8.- FANCG	75
1.3.5.9.- FANCI	76

1.3.5.10.- FANCJ	77
1.3.5.11.- FANCL	78
1.3.5.12.- FANCM	78
1.3.5.13.- FANCN	79
1.3.5.14.- FANCO	79
1.3.5.15.- FANCP	80
1.3.5.16.- FAN1	80
1.3.6.- Reparación de cruzamientos intercatenarios	81
1.3.7.- Anemia de Fanconi y actividad de LINE-1	88

2.- HIPÓTESIS Y OBJETIVOS 90

2.1.- Hipótesis 90

2.2.- Objetivos 90

2.2.1.- Determinar si se produce la retrotransposición por la vía ENi en células deficientes en la ruta de la Anemia de Fanconi 90

2.2.2.- Determinar las características de las inserciones de L1 producidas en células deficientes en la ruta de Anemia de Fanconi 90

2.2.3.- Establecer la relación entre la actividad de L1 y la inestabilidad genómica presente en la Anemia de Fanconi 90

2.2.4.- Determinar los niveles de retrotransposición en células deficientes en la ruta de la Anemia de Fanconi 90

MATERIAL Y MÉTODOS

3.- MATERIALES 92

3.1.- Antibióticos y productos químicos 92

3.2.- Medios de cultivo bacteriano 95

3.3.- Medios de cultivo celular 96

3.4.- Reactivos de transfección lipídicos 97

3.5.- Soluciones 98

3.6.- Cepas bacterianas 99

3.7.- Líneas celulares 100

3.7.1.- Humanas 100

3.7.2.- Hámster 102

3.8.- Vectores y plásmidos recombinantes 105

3.9.- RNAs de interferencia 108

3.10.- Oligonucleótidos sintéticos 112

4.- MÉTODOS	118
4.1.- Cultivos celulares	118
4.2.- Preparación de bacterias competentes	119
4.3.- Transformación de plásmidos superenrollados	121
4.4.- Aislamiento y purificación de DNA plasmídico	122
4.5.- Aislamiento y purificación de DNA genómico	123
4.6.- Cuantificación de ácidos nucleicos	124
4.7.- Transfección de vectores	127
4.8.- Preparación y transfección de RNAs de interferencia	128
4.9.- Cotransfección de vectores y RNAs de interferencia	131
4.10.- Ensayos en líneas celulares	133
4.10.1.- Ensayo de resistencia a mitomicina C	133
4.10.2.- Ensayo de retrotransposición	135
4.10.2.1.- Selección por neomicina	139
4.10.2.2.- Selección por blasticidina	140
4.10.2.3.- Con RNAs de interferencia	141
4.11.- Recuperación de inserciones de LINE-1	143
4.12.- Caracterización de las inserciones recuperadas	151

RESULTADOS

5.- RESULTADOS	154
5.1- ensayos de retrotransposición en células de hámster	154
5.1.1.- Ensayos de retrotransposición en células deficientes en FANCG	155
5.1.2.- Ensayos de retrotransposición en células deficientes en FANCC	157
5.1.3.- Ensayos de retrotransposición en células deficientes en FANCA	158
5.1.4.- Ensayos de retrotransposición en células deficientes en FANCD1/BRCA2	160
5.2.- Ensayos de retrotransposición en células de hámster complementadas	161
5.2.1.- Ensayos de retrotransposición en células CLV5B complementadas con el gen humano FANCC	162

5.2.2.- Ensayos de retrotransposición en células VH4 complementadas con el gen humano FANCA	164
5.3.- Ensayos de retrotransposición en fibroblastos de pacientes de Anemia de Fanconi	167
5.3.1.- Ensayos de retrotransposición en fibroblastos de pacientes deficientes en FANCG o FANCD1/BRCA2	167
5.3.2.- Ensayos de retrotransposición en fibroblastos de pacientes deficientes en FANCI o FANCD2	168
5.4.- Ensayos de retrotransposición en fibroblastos de pacientes de Anemia de Fanconi complementados	170
5.4.1.- Ensayos de retrotransposición en fibroblastos PD20F complementados con el gen FANCD2	170
5.4.2.- Ensayos de retrotransposición en fibroblastos ANKN complementados con el gen FANCI	173
5.5.- Recuperación y caracterización de eventos de retrotransposición	176
5.5.1.- Eventos de retrotransposición recuperados en células de hámster	178
5.5.1.1.- Inserciones recuperadas de elementos mutantes en el dominio endonucleasa (D205A)	178
5.5.1.2.- Inserciones recuperadas de elementos salvajes (WT)	179
5.5.2.- Eventos de retrotransposición recuperados en fibroblastos humanos	181
5.5.2.1.- Inserciones recuperadas de elementos mutantes en el dominio endonucleasa (D205A)	181
5.5.2.2.- Inserciones recuperadas de elementos salvajes (WT)	183
5.5.3.- Análisis de las inserciones anómalas recuperadas en líneas celulares deficientes en la ruta FA	187
5.5.3.1.- Inserciones con TSDs inusualmente largas	187
5.5.3.2.- Inserciones asociadas a deleciones en el sitio de inserción	188
5.5.3.3.- Inserciones con secuencias de relleno	189
5.6.- ensayos de retrotransposición en células tratadas con RNAs de interferencia	191

DISCUSIÓN

6.- DISCUSIÓN	201
6.1.- Actividad desregulada de L1 en la Anemia de Fanconi	202
6.1.1.- La retrotransposición de L1 por la vía ENi está aumentada en líneas celulares de hámster y fibroblastos de pacientes con defectos en la vía de reparación FA	203
6.1.2.- La complementación de líneas deficientes en miembros de FA revierte el aumento de movilidad de los elementos L1 mutantes en distintos miembros de la vía de FA	204
6.1.3.- Las inserciones de L1 en líneas deficientes en la ruta FA presentan características anómalas asociadas a procesos de inestabilidad genómica	205
6.1.4.- El silenciamiento de genes de la ruta FA y genes asociados altera los niveles de retrotransposición de L1	207
6.2.- Inestabilidad genómica causada por la actividad de L1 en la Anemia de Fanconi	209
6.3.- Vía de retrotransposición independiente de endonucleasa de L1 ante deficiencias en la reparación del DNA	211
6.4.- Reparación, replicación del DNA y actividad de L1	213
6.4.1.- Maquinaria de LINE-1 y la reparación del DNA en FA	215
6.4.2.- Recombinación homóloga acoplada a la replicación y actividad de L1 en FA	218
6.4.3.- Replicación del DNA y actividad de L1	226
6.4.4.- Deleciones y duplicaciones largas en el sitio de inserción: mecanismos moleculares	230
6.5.- Posibles implicaciones de la actividad de L1 en la progresión de la FA y la oncogénesis	235
6.6.- Posibilidad de tratamiento expandido de la FA	237

CONCLUSIONES

7.- CONCLUSIONES

242

BIBLIOGRAFÍA

8.- BIBLIOGRAFIA

244

INTRODUCCIÓN Y OBJETIVOS

1.- INTRODUCCIÓN

1.1.- Genomas y elementos móviles

1.1.1.- Complejidad de los genomas.

Charles Darwin y Alfred Wallace, en el año 1859, fue el primero en descifrar algunas de las claves de la evolución biológica y en describir un mecanismo que explicara el enorme potencial y la diversidad de las formas de vida (Wallace & Smith, 1858; Darwin, 1859). Desde nuestro punto de vista del siglo XXI, más actualizado, estos mecanismos ya no suponen ninguna sorpresa, pero resulta innegable que Darwin pudo llegar a esa conclusión sin conocer la existencia de la información genética en la manera en que la conocemos actualmente, aunque de alguna forma fue capaz de intuirlo.

Darwin, mediante la mera observación, llegó a deducir que la evolución es la consecuencia de una combinación de la herencia y la aleatoriedad del cambio heredable. Esto le permitió formular su “Teoría de la selección natural”, que sería la fuerza que gradualmente va adaptando a los seres vivos a las condiciones ambientales. No obstante, los mecanismos biológicos que se ocultaban detrás de las observaciones de Darwin seguirían siendo un misterio durante varias décadas.

Años más tarde, tras la experimentación en *Drosophila* llevada a cabo por el grupo del Dr. Morgan, tuvo lugar la primera propuesta de la existencia de un genoma (Morgan, 1926). Posteriormente se produjo el descubrimiento de la

estructura del DNA y el código genético (Griffith, 1928; Avery *et al.*, 1944; Watson & Crick, 1953; Gardner *et al.*, 1962; Wahba *et al.*, 1963), y se retomó la visión de Darwin, estableciéndose los mecanismos de herencia y mutación. Ahora bien, después de establecer el concepto de genoma como un principio, se planteó un enigma de difícil respuesta: ¿es el genoma una serie de bloques moldeado por la selección natural, se trata más bien de una combinación aleatoria de genes, o tal vez una combinación de ambos conceptos?

Nuestro conocimiento de los genomas empezó a aumentar en el año 1976, cuando se secuenció por primera vez un genoma completo, el del bacteriófago MS2 (Fiers *et al.*, 1976); desde entonces se han ido secuenciado multitud de genomas de los diferentes reinos de seres vivos, cambiando por completo nuestro entendimiento acerca de los genomas, de su estructura, y de las relaciones entre ambos. La secuenciación del genoma humano se completó en el año 2001 (Lander *et al.*, 2001), y nos permitió descubrir datos tan sorprendentes como que tan sólo un 1-2% de nuestro DNA codifica para genes, mientras que el resto no es codificante. Unos años más tarde, tras completarse la secuenciación del genoma de chimpancé, se pudo comprobar que éste presenta una homología del 96-98% con el genoma humano (Newman *et al.*, 2005).

Sin embargo, mucho antes de que concluyese la secuenciación del primer genoma, se produjo una de las mayores sorpresas que han tenido lugar a lo largo de muchos años de investigación en el campo de la Genética. La investigadora Barbara McClintock describió la presencia de DNA móvil en la planta del

maíz (McClintock, 1950). Hasta ese momento se pensaba que el DNA debía de ser algo estático y estable, inmóvil, con una tasa de mutación estimada de 10^{-8} por nucleótido, un número realmente bajo, que respaldaba el concepto de que el genoma debía ser algo estático. Este descubrimiento plantearía a su vez muchas preguntas como: ¿Cómo y por qué existen los elementos móviles? ¿Tienen alguna relación con la evolución? Y si nos centramos exclusivamente en el individuo ¿Qué efectos pueden tener en el desarrollo de su vida? ¿Pueden afectar a los descendientes del individuo? ¿Cuándo tiene lugar su movilización? Las respuestas a muchas estas preguntas aún no están claras, a pesar de que cada vez se conocemos más acerca de estos elementos.

1.1.2.- Elementos móviles.

Los elementos móviles son secuencias repetidas que poseen la capacidad, de moverse, como su propio nombre indica, dentro de los genomas de las células que los portan o “genomas hospedadores” (Doolittle & Sapienza, 1980; Orgel & Crick, 1980).

De acuerdo con la teoría de Darwin, se especuló que la presencia de los elementos móviles en los genomas podía aportar alguna ventaja a los organismos (Blot, 1994). Anteriormente, en los años 80, se propuso que la expansión de estos elementos podía deberse únicamente a su capacidad para replicarse dentro de los genomas (Orgel & Crick 1980). Su capacidad de replicarse les proporciona una ventaja con respecto a otros genes a la hora de transmitirse de una generación a otra, lo que hizo que

se empezaran a considerar como DNA egoísta y parásito, que carecía de utilidad para el hospedador y de importancia en su evolución. En definitiva, fueron considerados como DNA “basura” (“*junk DNA*”) (Orgel & Crick 1980).

No obstante, poco a poco fueron apareciendo nuevas evidencias a nivel molecular mostrando que, a pesar de que en su mayoría pueden considerarse como auténticos parásitos del genoma, estos elementos podían, en algunos casos, proporcionar características que sirvieran a los intereses de sus hospedadores. Por ejemplo, en genomas de bacterias donde parecen cooperar para la coevolución de ambos (Blot, 1994), o en *Saccharomyces cerevisiae*, donde se han encontrado asociados a reordenaciones del genoma favorables para la evolución de la especie (Dunham, 2002).

Si tenemos en cuenta que la evolución se basa en la generación de la novedad o variedad a partir del material preexistente, más que en la aparición espontánea de características ausentes anteriormente, los elementos móviles constituyen una base perfecta para la modificación. En algunos casos, ésta puede llevar a la obtención de características útiles para sus hospedadores (McDonald, 1993); y de hecho, se han descrito varios ejemplos de este fenómeno (Wessler *et al.* 1995, Britten 1996), que se conoce como “domesticación molecular” (Miller *et al.*, 1997). Este fenómeno consiste en que el genoma hospedador utiliza parte de la secuencia del elemento móvil, para construir una nueva función, como por ejemplo ocurre con un elemento P en *Drosophila* que ha dado lugar a una proteína con capacidad represora de los propios

elementos móviles, protegiendo así el genoma de la especie (Nouaud & Anxolabéhère, 1997).

Por tanto, a lo largo de millones de años de evolución, los seres vivos han tenido que tratar con estos elementos, controlando su expansión en sus genomas; pero a su vez, de manera accidental o motivados por la propia selección natural, han conseguido domesticarlos y usarlos para su propio beneficio. Estos elementos, debido a su movilidad, son en sí mismos una fuente potencial de variabilidad, y por lo tanto, contrariamente a lo que se pensaba en los años 80, pueden actuar como verdaderos motores evolutivos (revisado en Kazazian, 2004).

En muchas ocasiones, como por ejemplo ocurre con el retrotransposón humano LINE-1, estos elementos contienen secuencias reguladoras, tales como promotores, sitios de unión de factores de transcripción o islas CpG (Swergold, 1990; Speek *et al.*, 2001; Yang *et al.*, 2003; Athanikar *et al.*, 2004). Esto los hace especialmente interesantes a nivel de regulación génica, ya que no sólo son capaces de afectar a los genomas de los individuos a nivel de secuencia de DNA (Kazazian, 1999), sino que también pueden afectar a el modo en el que se expresan los genes, o incluso a el estado de la cromatina en el lugar donde se movilizan (García-Pérez *et al.*, 2010).

Todas estas características han hecho de los elementos móviles unos aliados valiosos para la evolución de los seres vivos como ya se ha mencionado; sin embargo, son estas mismas capacidades las que los convierten en un riesgo para los individuos, ya que su movilización en áreas críticas para el funcionamiento correcto de la célula pueden conducir a la muerte de ésta, o a la pérdida de

capacidades necesarias para el desarrollo adecuado del individuo. Precisamente es este último aspecto el que da una gran relevancia a su estudio en el campo de la Biomedicina, ya que la actividad de estos elementos puede ser la causa de una enfermedad genética. En humanos, el primero que describió este fenómeno fue el Dr. Haig Kazazian en el año 1988; encontró que una inserción del retrotransposón LINE-1 en un exón del gen del factor VIII era causante de hemofilia A (Kazazian, 1988). Otro ejemplo de patología humana que puede ser causada por la inserción de un retrotransposón lo encontramos en la distrofia muscular de Duchenne (Miciak, 1992). Desde entonces, se han caracterizado cerca de 100 enfermedades genéticas en seres humanos producidas por la movilidad de LINE-1 (Hancks *et al.*, 2012).

1.1.3.- Clases de elementos móviles.

Existen muchos tipos de secuencias móviles, que se pueden englobar dentro de 2 grandes grupos:

- Elementos que utilizan un **intermediario de DNA para movilizarse** (o elementos móviles de clase II) y,

- Elementos que utilizan un **intermediario de RNA durante su movilización** (elementos móviles de clase I).

La variedad dentro de cada uno de estos grupos es muy amplia (revisado en Goodier and Kazazian, 2008), y el hecho de que se encuentren presentes en casi todos los seres vivos, unido a que en cada grupo evolucionen de manera independiente,

complica enormemente su clasificación, pero a grandes rasgos los agrupamos en la siguiente tabla:

Elemento móvil	Organismo	Intermediario
Elementos IS (secuencias de inserción bacterianas)	Procariotas	DNA
Transposones bacterianos	Procariotas	DNA
Transposones autónomos de DNA subclase I (de cortar-pegar)	Eucariotas	DNA
Polintones/Mavericks (subclase II de transposones de DNA)	Eucariotas y Procariotas	DNA
Helitrones (subclase II de transposones de DNA)	Eucariotas y Procariotas	DNA
Transposones de DNA no autónomos	Eucariotas y Procariotas	DNA
Intrones tipo II	Procariotas	RNA
Retrotransposones LTR	Eucariotas	RNA
Retrotransposones LTR no autónomos	Eucariotas	RNA
Retrotransposones no LTR autónomos	Eucariotas	RNA
Retrotransposones no LTR no autónomos	Eucariotas	RNA

Tabla 1: Tipos de elementos móviles. Elementos móviles presentes en los seres vivos. Abreviaturas: LTR (*Long Terminal Repeats*): terminaciones largas repetidas.

Sin embargo, aunque los elementos móviles están presentes en casi todos los seres vivos, tanto el tipo de elemento como su proporción en los genomas varían enormemente de una especie a otra (figura 1). Por ejemplo, en el genoma de *Caenorhabditis elegans* el 90% de los elementos móviles pertenecen a DNA transposones, en *S. cerevisiae* sólo aparece un tipo de retrotransposón, mientras que en humanos un 75%

de los elementos móviles pertenece a un solo tipo de retrotransposón (revisado en Feschotte and Pritham, 2007).

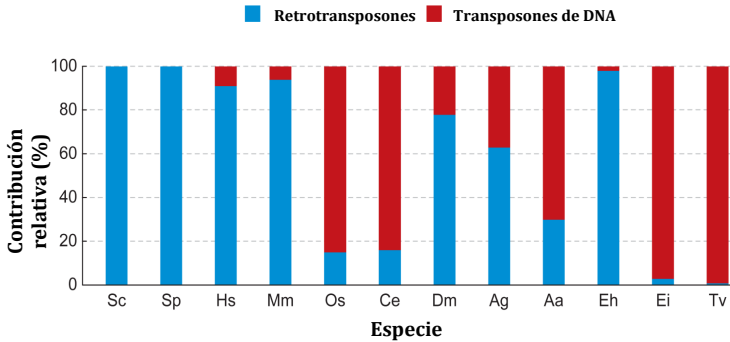


Figura 1: Proporción relativa de elementos móviles en genomas de distintas especies. La figura muestra el porcentaje que representa cada clase de elemento móvil (retrotransposones en rojo y transposones de DNA en azul). Abreviaturas de nombres de especies: *Sc*: *Saccharomyces cerevisiae*; *Sp*: *Schizosaccharomyces pombe*; *Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Os*: *Oryza sativa*; *Ce*: *Caenorhabditis elegans*; *Dm*: *Drosophila melanogaster*; *Ag*: *Anopheles gambiae*; *Aa*: *Aedes aegypti*; *Eh*: *Entamoeba histolytica*; *Ei*: *Entamoeba invadens*; *Tv*: *Trichomonas vaginalis*. (Adaptado de Feschotte & Pritham, 2007).

1.1.3.1.- DNA Transposones.

Los DNA transposones son elementos móviles que se caracterizan por dispersarse por el genoma utilizando un intermediario de DNA de cadena sencilla (descrito únicamente en procariontes (Ton Hoang *et al.*, 2005) o de cadena doble; dicho intermediario se inserta en una nueva localización del

genoma, generalmente mediante la acción de una transposasa (Craig *et al.*, 2002; Ton-Hoang *et al.*, 2005). Están presentes en casi todos los seres vivos (Figura 2, Feschotte & Pritham, 2007), aunque en algunos, como ocurre en el caso de los seres humanos, han perdido su capacidad para seguir movilizándose.

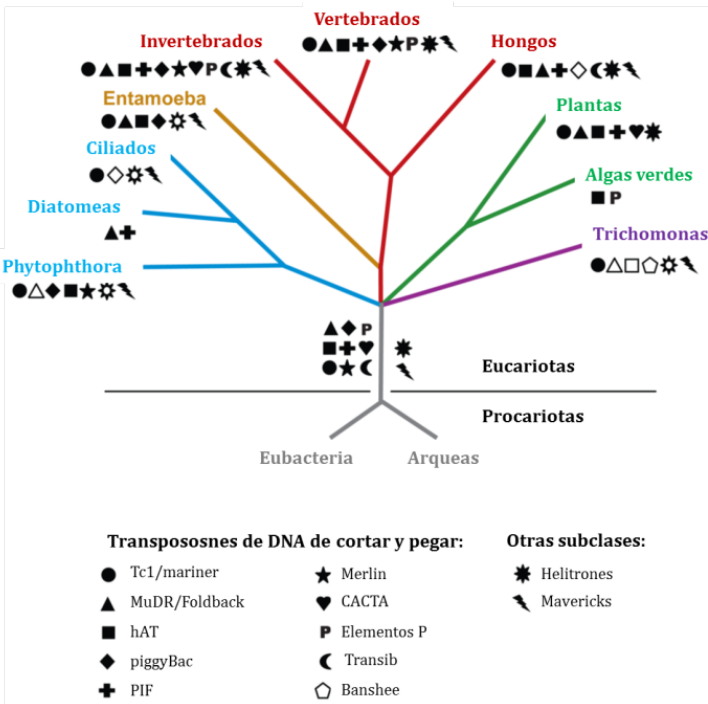


Figura 2: Transposones de DNA en los seres vivos
 Representación de los diferentes tipos de transposones de DNA presentes en cada uno de los grupos de seres vivos (Adaptado de Feschotte & Pritham, 2007).

TRANSPOSONES CON INTERMEDIARIO DE DNA

Tipo	Características estructurales	Mecanismo de movilización	Ejemplos
Elementos IS	Repeticiones invertidas de ≈50pb que flanquean una transposasa y, en algunos, una resolvasa	Escisión o copia del DNA e integración en el sitio diana	IS1, IS10
Transposones bacterianos	Gen de resistencia a antibiótico en la región central flanqueado por elementos IS	Copia del DNA e integración en el sitio diana	Tn9
Transposones de cortar y pegar	TIRs que flanquean una región que codifica una transposasa	Escisión del DNA e integración en el sitio diana mediante la transposasa	Elementos P (<i>Drosophila</i>)
Polintons o Mavericks	Largas TIRs, codifican una polimerasa y una integrasa.	Parecen movilizarse utilizando una polimerasa codificada por ellos mismos	Polinton-1_DR (<i>Danio Rerio</i>)
Helitrones	No poseen TIRs, pero sí un dominio replicasa y un dominio helicasa.	Relacionado con la replicación del círculo rodante, autónomos o no autónomos	HeliBat1 (<i>Myotis lucifugus</i>)
MITES	Secuencias de 80-600pb, que contienen las TIRs pero no codifican una transposasa	Escisión del DNA e integración en el sitio DNA mediante la transposasa de un transposón autónomo	Stowaway

Tabla 2: Tipos de transposones de DNA. La tabla muestra los principales tipos de elementos de DNA móvil presentes en los seres vivos. Abreviaturas: IS: *Insertional Sequence*; TIR: *Terminal Inverted Repeats*.

A pesar de que podemos encontrarlos tanto en procariontas como en eucariontas (tabla 2), sus características y mecanismo de movilización varían, pero centrándonos exclusivamente en los elementos móviles de DNA de eucariontas podemos distinguir:

- **Transposones de DNA de cortar y pegar:** estos elementos se escinden como DNA de doble cadena y se reintegran en otra parte del genoma; deben su nombre a que abandonan su lugar original para de su replicación (Kapitonov & Jurka, 2006; Pritham *et al.* 2007), aunque su mecanismo de movilización aún no

se ha descrito completamente. También forman parte de la subclase II de transposones de DNA.

- **MITEs** (*Miniature Inverted repeats Transposable Elements*): en este grupo encontramos transposones de DNA no autónomos, que requieren la intervención de la transposasa codificada por los transposones autónomos. Son elementos muy diversos, pero generalmente cortos, de unas 80-600 pares de bases de longitud, que poseen las TIRs características de los transposones autónomos (o similares), pero que no codifican su propia transposasa.

Los DNA transposones mejor conocidos pertenecen mayoritariamente al grupo de los transposones de cortar y pegar. Se movilizan abandonando su sitio en el genoma para ocupar uno nuevo (Engels *et al.*, 1990). Típicamente las terminaciones invertidas terminales o TIRs flanquean la secuencia que codifica la transposasa (Craig *et al.*, 2002).

La transposasa reconoce y corta las TIRs, liberando así la secuencia completa del transposón; posteriormente, para la mayoría de las transposasas, sostiene los extremos del transposón juntos mientras encuentra otra localización genómica donde integrarlo (Mizuuchi, 1992). El sitio de inserción suele ser aleatorio, aunque es muy frecuente que sea un sitio cercano a la localización genómica de la que se escinde el elemento, conociéndose este fenómeno como *local hopping* (Bancroft, 1993).

Dado que un DNA transposón no suele dejar una copia en el sitio del que se ha liberado, este mecanismo se llama de “cortar y pegar” (Craig *et al.*,

2002). Aunque a veces pueden generarse dos copias del transposón en la nueva localización genómica (Delattre *et al.*, 2000), la manera de propagarse por los genomas puede producirse por dos mecanismos:

- Mediante la recombinación homóloga: al escindirse el transposón de una cadena de DNA genera una rotura de doble cadena, ésta puede ser reconocida por la maquinaria de reparación celular y reparada mediante la recombinación homóloga, por lo que se puede volver a introducir la copia del elemento presente en el otro cromosoma (Engels *et al.*, 1990).

- Una segunda opción descrita recientemente horquillas de replicación (Spradling *et al.*, 2011), de esta manera pueden ser duplicados por las replicasas del hospedador antes de movilizarse a una nueva localización o aterrizar en una región justo antes de que sea replicada.

- Una combinación de los dos mecanismos anteriores supone un escenario idóneo para el elemento, que puede escindirse generando una rotura de doble cadena en su localización original a la vez que se introduce en una región que va a ser replicada (Spradling *et al.*, 2011). La copia que aterriza en una horquilla de replicación se replica utilizando la maquinaria celular y se establece en ambos cromosomas, mientras que en la localización original la rotura de doble cadena se repara mediante la recombinación homóloga, regenerando la copia perdida. De esta manera el elemento puede ganar hasta 2 copias nuevas netas en el proceso completo. Todo el proceso está favorecido durante la fase S, ya que las roturas de doble cadena se reparan preferentemente mediante la recombinación

homóloga durante la replicación del DNA (Deans & West, 2011).

A pesar de todo lo mencionado anteriormente, los transposones de DNA pierden de manera relativamente rápida su actividad en los genomas, debido a mutaciones de distintos tipo que afectan a la secuencia de la transposasa, dejándola inactiva, pero que no afectan a las TIRs. Así, cada vez hay menos transposasas activas para reconocer más TIRs, y la secuencia de la transposasa activa puede llegar incluso a desaparecer (Hart *et al.*, 1997). Para evitar su extinción, la mejor opción para el elemento reside en la transferencia horizontal (Hart *et al.*, 1997). Esto ocurre normalmente entre especies cercanas evolutivamente, como por ejemplo se ha descrito en *Drosophila* (Loreto *et al.*, 2008). En la transferencia horizontal el elemento pasa de su hospedador original a uno nuevo, invadiendo su genoma, que no posee mecanismos adecuados para controlar su movilidad y no presenta muchas secuencias como las TIRs propias del elemento, lo que puede dar lugar a una expansión inicial muy rápida del transposón en el genoma de su nuevo hospedador (Hart *et al.*, 1997; Loreto *et al.*, 2008).

1.1.3.2 Retrotransposones.

El segundo gran grupo principal de elementos móviles son los llamados retrotransposones o retroelementos. Estos elementos tienen en común que durante su proceso de movilización utilizan un RNA intermediario que, mediante una transcriptasa inversa (o reversotranscriptasa), se retrotranscribe a una copia de DNA que se inserta en el genoma. A diferencia de lo que ocurre con los transposones de DNA que utilizan el mecanismo “cortar y pegar” o “no replicativo”, los retrotransposones no abandonan en ningún momento su localización inicial, por lo que su mecanismo de movilización se llama de “copiar y pegar” o “replicativo” (Kazazian & Scott, 1993; Luan *et al.*, 1993; revisado en Goodier & Kazazian, 2008; Havecker *et al.*, 2004 y Malik & Eickbush, 2001).

Al igual que ocurre con los transposones de DNA, existen muchos tipos de retrotransposones, tanto autónomos como no autónomos (tabla3), y podemos encontrarlos en casi todos los grupos de seres vivos; si bien, existe una diferencia importante entre ellos en el momento y el lugar en el que se produce la transcripción inversa de su RNA a DNA. La clasificación de los retrotransposones se realiza en base a sus características estructurales, que determinan su mecanismo de movilización (tabla 3, figuras 3 y 4).

RETROTRANSPOSONES (INTERMEDIARIO DE RNA)

Tipo	Características estructurales	Mecanismo de movilización	Ejemplos
Intrones tipo II (tipo no-LTR)	RNA catalítico altamente estructurado, codifica una IEP con actividad RT, matorasa y normalmente una EN	Retrohoming: formación de una RNP que reversotranscribe el RNA a DNA en el sitio diana	Ecl5 (<i>E.coli</i>), L1.LtrB (<i>Lactococcus</i>)
LTR	LTRs de ≈250 a 600pb, que flanquean una región que contiene una RT, una integrasa y una proteína similar a las "Gag" de retrovirus, y en algunos, un gen env	Transcripción del RNA desde la LTR de 5' del elemento, formación de una VLP, retrotranscripción por la RT a DNA y posterior integración en el sitio diana	Ty1, Ty3 (<i>Drosophila</i>), ERVs (mamíferos)
LTR no autónomos	LTRs de ≈250 a 600pb, que flanquean una región que puede contener o no proteínas típicas de los autónomos, pero no todas	Transcripción del RNA desde la LTR de 5' del elemento, capturan la maquinaria de los elementos LTR autónomos	Dasheng, Zeon-1 (maíz)
No-LTR autónomos	Longitud variable, generalmente largos, actividad promotora en su región 5', codifican una RT y normalmente una EN, contienen una región rica en A/T en 3'	Transcripción del elemento desde la región 5', formación de una RNP, integración mediante TPRT: reversotranscripción del RNA a DNA en el sitio de inserción	Elementos F y G (<i>Drosophila</i>); R2Bm (<i>Bombyx mori</i>), LINES
No-LTR No autónomos	Longitud variable, generalmente cortos, no codifican RT, contienen una región rica en A/T en 3'	Transcripción del elemento, su RNA captura la maquinaria de los elementos no-LTR autónomos	SINES (mamíferos); pseudogenes procesados

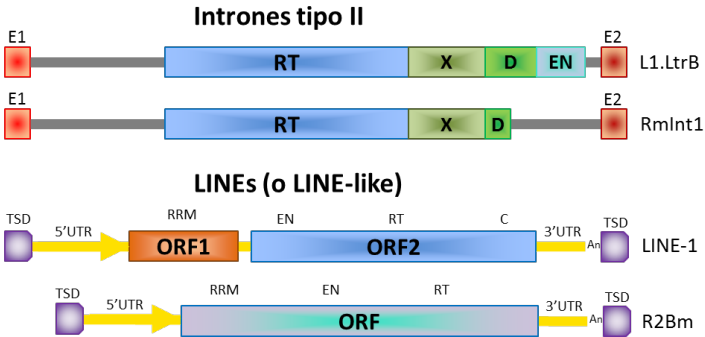
Tabla 3: Tipos de retrotransposones. Abreviaturas: IEP (*Intron Encoded Protein*), proteína codificada por el intrón; RT, reversotranscriptasa; EN, endonucleasa; VLP (*Viral-Like Particle*), partícula similar a las partículas virales; RNP (*RiboNucleic Particle*), partícula de RNA y proteína(s); TPRT (*Target-Primed Reverse Transcription*), reversotranscripción iniciada en el sitio diana; ERVs (*Endogenous RetroVirus*), retrovirus endógenos

Existen dos grandes tipos de retrotransposones atendiendo a la presencia o ausencia de las LTRs. Así, los elementos que las poseen se llaman retrotransposones tipo LTR y los que carecen de ellas retrotransposones tipo no LTR (revisado en Goodier & Kazazian, 2008). Los intrones

tipo II son retroelementos de procariotas de tipo no-LTR, pero se consideran como grupo independiente por presentar grandes diferencias con respecto al resto de retrotransposones no-LTR (Lambowitz & Zimmerly, 2011).

Si nos centramos en los retroelementos autónomos, responsables de su movilidad y de la movilidad de los elementos no autónomos, encontramos características similares dentro de los elementos del mismo tipo (figura 3); sin embargo, entre los diferentes tipos de retroelementos autónomos, la única característica que comparten es la presencia de un dominio reversotranscriptasa (o RT) funcional.

TIPO NO LTR



TIPO LTR

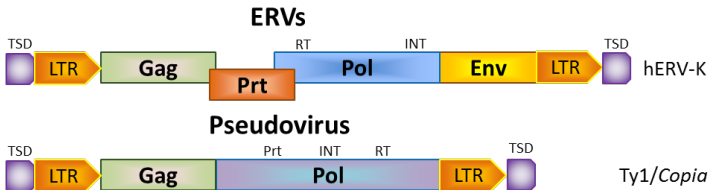


Figura 3: Retrotransposones autónomos. LTR (*Long Terminal Repeat*), terminación larga repetida directa. **Intrones tipo II:** E1, exón 1; E2, exón 2; RT, reversotranscriptasa; X, dominio maturasa; D, dominio de unión a DNA; EN, dominio endonucleasa. Ejemplos: L1.LtrB de *Lactococcus*, Rmlnt1 de *Sinorhizobium*. **LINEs** (*Long Interspersed Elements*): TSD (*Target Site Duplication*), duplicación del sitio diana; UTR (*UnTranslated Region*), región no traducida; RRM (*RNA Recognition Motif*), dominio de unión a RNA; EN, endonucleasa; RT, reversotranscriptasa; C, dominio rico en cisteínas; An, cola de poliA. Ejemplos: LINE-1 de mamíferos, R2Bm de *Bombyx mori*. **ERVs y pseudovirus** (*Endogenous RetroVirus*): TSD, duplicación del sitio diana; Prt, proteasa; RT, retrotranscriptasa; INT, integrasa; Env (*Envelope*), proteína de la cubierta. Ejemplos: hERV-K de mamíferos, Ty1/*Copia* de *Drosophila*. (Adaptado de Lambowitz & Zimmerly, 2011; Goodier & Kazazian, 2008; Mills *et al.*, 2007; Malik & Eickbush, 2000; Havecker *et al.*, 2004)

1.3.2.1.- Intrones tipo II.

Estos elementos están presentes en los genomas de procariotas, arqueas y en los orgánulos celulares de algunos eucariotas. Se pueden clasificar como retroelementos no-LTR al carecer de dichas repeticiones, pero presentan algunas diferencias con los retrotransposones no-LTR de eucariotas.

Los intrones tipo II codifican un RNA catalítico altamente estructurado que presenta 6 dominios estructurales, denominados de DI a DVI (figura 4). En su dominio IV, codifican una proteína llamada IEP (*Intron Encoded Protein*), que contiene la actividad reversotranscriptasa (RT) y maturasa, un motivo de unión a DNA, y generalmente un dominio endonucleasa. La IEP y el RNA conforman una partícula ribonucleoproteica (RNP, por *RiboNucleoproteic Particle*) (revisado en Lambowitz & Zimmerly, 2011).

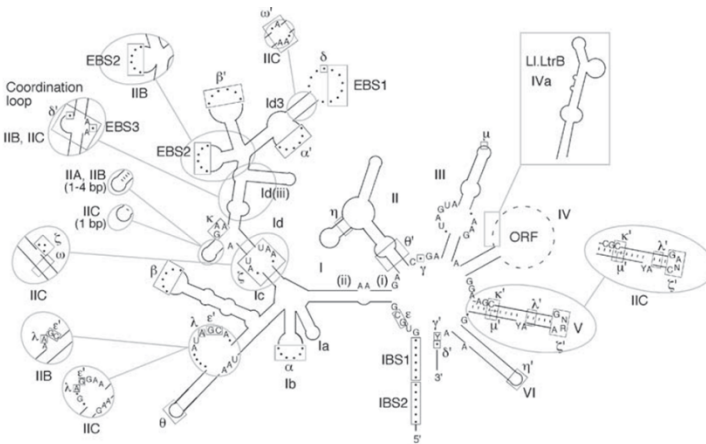


Figura 4: Estructura secundaria del RNA de un intrón tipo II. (Adaptado de Lambowitz & Zimmerly, 2011).

El elemento completo es transcrito a RNA, que es liberado por *splicing* y madura mediante su actividad catalítica, adquiriendo su conformación estructural secundaria. La IEP se puede unir entonces al RNA conformando la RNP activa. Esta RNP reconoce una región específica en el genoma bacteriano y produce un corte en una de las cadenas del DNA, tras lo cual inicia su retrotranscripción o *splicing* reverso, mediante el cual genera una cadena de DNA copia del elemento (Belfort *et al.*, 2002, Mobile DNA II). La síntesis de la cadena complementaria y los procesos finales de la integración parece que son realizados por la maquinaria del hospedador (Yao *et al.*, 2013).

A este mecanismo se le conoce como *retrohoming* (figura 5), y es similar en muchos aspectos a la reacción “TPRT” de los elementos no LTR de eucariotas (descrita más adelante). Por último, cabe destacar que aunque los intrones tipo II normalmente reconocen una secuencia consenso, estos elementos pueden integrarse en sitios distintos, conociéndose este fenómeno como “*retrohoming* ectópico” o retrotransposición, aunque esto es mucho menos frecuente (figura 5) (Belfort *et al.*, 2002, Mobile DNA II).

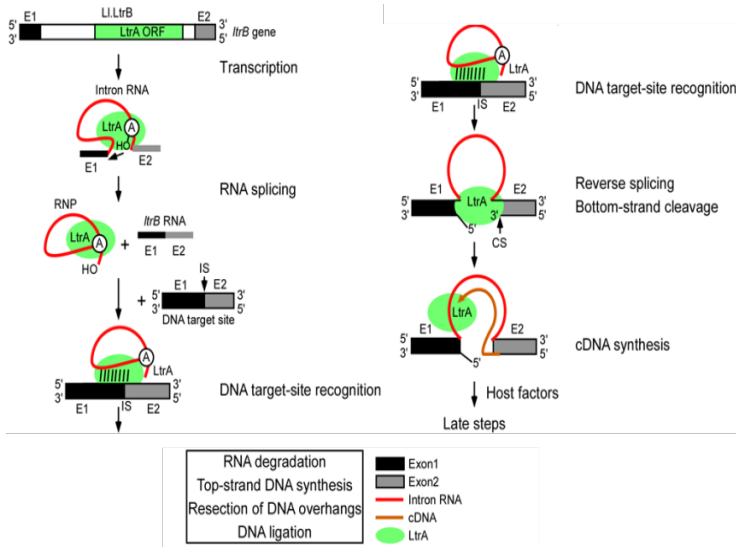


Figura 5: Retrohoming: mecanismo de movilización de un intrón tipo II. (Adaptado de Yao *et al.*, 2013)

1.1.3.2.2.- Retrotransposones LTR.

Este tipo de retrotransposones deben su nombre a la presencia de terminaciones largas directas repetidas (*Long Terminal Repeats* o LTRs) (Kennedy *et al.*, 1982), a diferencia de las TIRs de los transposones de DNA, estas LTRs se encuentran en el mismo sentido 5'-3' al principio y al final del elemento. Estas repeticiones varían de tamaño según el elemento, aunque suelen ser de unas 300-1000 pares de bases (Kennedy *et al.*, 1982). En general, los retrotransposones tipo LTR son muy similares a los retrovirus en varios aspectos de su biología.

Estos retrotransposones contienen un promotor en la región 5' de su LTR que dirige la transcripción del elemento. Una vez el rna se encuentra en el citoplasma, las proteínas codificadas por el gen *gag* se encargan de empaquetar su propio RNA, formando unas vesículas muy similares a las de los virus, llamadas VLPs (*Viral-Like Particles*) (Pochart *et al.*, 1993). Dentro de estas partículas se retrotranscribe el RNA del elemento a DNA por la RT codificada por el gen *pol*, en un proceso complejo que se produce en varios pasos (Rachidi *et al.*, 2005). Posteriormente el DNA de doble cadena sale de las VLPs y vuelve al núcleo para integrarse en el genoma hospedador mediante la acción de una integrasa también codificada por el gen *pol* del elemento (Wilhelm & Wilhelm, 2001). Su mecanismo de movilización se resume de forma esquemática en la figura 6. Frecuentemente aparecen flanqueados por duplicaciones cortas del sitio diana de inserción, llamadas TSDs (de *Target Site Duplications*), que se generan en el proceso de integración (Sverdlov, 1998).

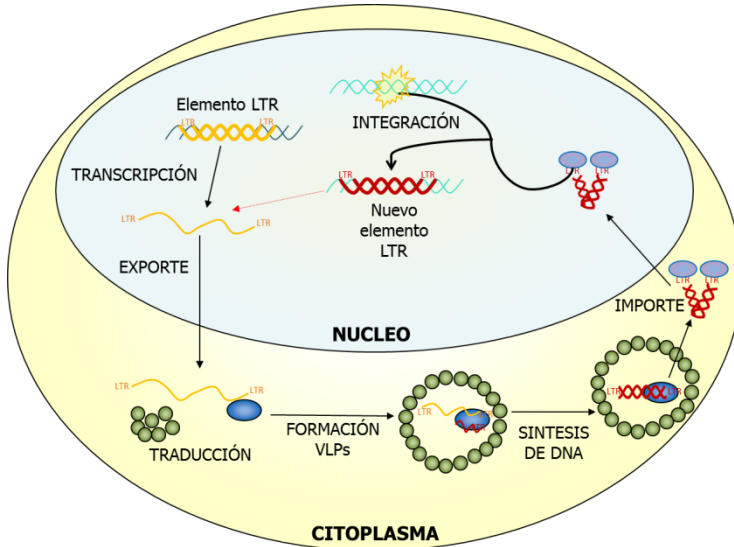


Figura 6: Mecanismo de movilización de un retrotransposón de tipo LTR. Los elementos LTR son transcritos y exportados al citoplasma, donde forman las VLPs, y sintetizan su DNA, que vuelve al núcleo para integrarse en el genoma.

El mecanismo de replicación de estos retrotransposones es casi idéntico al de los retrovirus, y ambos codifican proteínas muy similares (Wilhelm & Wilhelm, 2001); no obstante, los retrovirus parece que han evolucionado a partir de ellos, mediante la adquisición de una propiedad adicional, la capacidad de formar una cubierta que protege el material genético del virus en su viaje de una célula a otra, esto sería posible con la adquisición de un gen *env* por parte de un elemento LTR; otros genes accesorios asimismo podrían haber dado lugar a otros tipos de virus (Purcell *et al.*, 1996; Lerat & Capy, 1999). Como ejemplo, el virus VIH que causa la

enfermedad humana Síndrome de la Inmunodeficiencia Adquirida o SIDA, es un retrovirus complejo, con varios genes accesorios que permiten entre otros procesos la evasión del sistema inmunitario (Piguet *et al.*, 2000).

Se ha propuesto que este gen *env* puede, en algunos casos, convertir al elemento en infectivo, permitiendo el paso de una célula a otra, como ocurre con los elementos *gypsy* de *Drosophila* (Song *et al.*, 1994; Kim *et al.*, 1994).

La situación se complica cuando examinamos ciertos tipos de retroelementos LTR, que poseen un gen similar a los *env* de retrovirus; este es por ejemplo el caso de los retrovirus endógenos (ERV por *Endogenous Retroviruses*), estos retroelementos parecen haberse originado a partir de un retrovirus ancestral, que infectó la línea germinal y se integró de manera permanente en el genoma del hospedador (Yi *et al.*, 2004).

Por último, los **retrotransposones de tipo LTR no autónomos** son elementos que poseen unas LTRs idénticas o similares a las del elemento autónomo responsable de su movilización, pero carecen de alguno o todos los genes necesarios para su replicación e integración. La maquinaria de un retroelemento LTR autónomo puede reconocer estas LTRs y llevar a cabo el ciclo. Este es el caso por ejemplo del elemento no autónomo *Dasheng* del maíz y su correspondiente autónomo RIRE2 (Jiang *et al.*, 2002).

1.1.3.2.2.- Retrotransposones No-LTR.

Este tipo de retrotransposones deben su nombre a que carecen de las LTRs, y tanto su estructura como su mecanismo de movilización son muy diferentes a los retrotransposones de tipo LTR y los retrovirus. Se pueden dividir en dos grupos principales atendiendo a la longitud de su secuencia:

- Los **LINEs** (*Long INterspersed Elements*) o LINE-like, con una secuencia más larga (generalmente alrededor de 6kB), suelen ser retrotransposones autónomos (Revisado en Kazazian & Goodier, 2002).

- Los **SINEs** (*Short INterspersed Elements*), más cortos, generalmente de unos pocos cientos de bases, que son retrotransposones no autónomos (revisado en Okada, 1991).

Los LINEs suelen contener una secuencia promotora interna situada en su región 5' no traducida (5'UTR) para iniciar su expresión (Swergold, 1990), una región intermedia, de longitud y características variables según el tipo de elemento, que contiene las proteínas del elemento (ORFs); y por último, una secuencia no traducida en su extremo 3' (3'UTR) rica en adeninas y timinas que acaba en una señal de poliadenilación débil.

Los LINEs (figura 3) constituyen el grupo de retroelementos con mayor éxito evolutivo en mamíferos, siendo LINE-1 (figura 10), descrito con más detalle en el siguiente apartado, su representante más extendido. La región central de estos elementos móviles está compuesta por una secuencia que codifica dos proteínas necesarias para la retrotransposición (Kulpa & Moran, 2005),

denominadas como ORF1 y ORF2; sin embargo, en otros retroelementos de tipo no LTR, sólo se requiere una poliproteína, como ocurre en el caso de los elementos R2 de *Bombyx mori* (Luan *et al.*, 1993).

Para su movilización, el RNA resultante de la transcripción del elemento va al citoplasma (Skorowski & Singer, 1985), donde se traducen las proteínas necesarias para la retrotransposición (Alisch *et al.*, 2006; Dmitriev *et al.*, 2007) y se unen al RNA formando una ribonucleopartícula (RNP), que vuelve al núcleo (Kulpa & Moran, 2006). Una vez allí, se produce la reversotranscripción *in situ* e integración del elemento, en un proceso denominado *Target Primed Reverse Transcription* o TPRT (Luan *et al.*, 1993), descrito con más detalle más adelante (figura 11).

El sitio de inserción de estos elementos varía dependiendo del elemento, en algunos casos es muy específico, como ocurre por ejemplo con los elementos R1 y R2 de insectos, que sólo se insertan en secuencias muy concretas de los genes de RNA ribosomal (Luan *et al.*, 1993); en otros casos, en cambio, el sitio diana es mucho más variable, pudiendo insertarse en casi cualquier región del genoma, como ocurre por ejemplo con los elementos LINE-1 humanos (L1), que a nivel de secuencia sólo necesitan la presencia de una región corta rica en AT para iniciar su inserción por TPRT (Jurka, 1997).

Los SINEs y otros retroelementos no LTR no autónomos carecen de reversotranscriptasa, y por lo general su secuencia carece de potencial codificante, aunque poseen la capacidad de secuestrar de alguna manera la maquinaria de los retrotransposones autónomos (Dewannieux *et al.*, 2003; Hancks *et al.*,

2012). En este tipo de elementos apreciamos una región rica en adeninas y timinas en el extremo 3', por donde empiezan la reversotranscripción los retrotransposones no LTR autónomos.

A este grupo pertenecen retrotransposones como los elementos B1 de ratón o Alu o SVA (por *SINE VNTR Alu*) en humanos, cuya movilidad es dirigida por los elementos LINE (revisado en Babushok & Kazazian, 2007).

1.1.4.- Elementos móviles en el genoma humano.

Una de las grandes sorpresas tras terminar la secuenciación del genoma humano fue comprobar que, a pesar de su gran tamaño, tan sólo un 1-2% de su contenido correspondía a secuencias codificantes (Lander *et al.*, 2001). Se comprobó que una enorme proporción del genoma (más del 50%) estaba formada por diferentes tipos de secuencias repetidas, destacando entre ellas los elementos móviles, que suponen aproximadamente un 45% del genoma (Tabla 4 y Figura 7), destacando la importante presencia de los retroelementos de tipo no-LTR (Lander *et al.*, 2001; Cordaux & Batzer, 2009). Por otro lado, estudios más recientes indican que hasta más de un 70% del genoma humano es derivado de la actividad de DNA móvil, aunque la presencia de mutaciones dificulta su clasificación (de Koning *et al.*, 2011).

Si analizamos en mayor profundidad la proporción de elementos móviles en el genoma humano encontramos la siguiente distribución, recogida en la tabla 4 y la figura 7.

Elementos	Fracción de genoma (%)	Específicos de primate (%)
LINEs	20.99	7.94
LINE1	17.37	7.94
LINE2	3.30	-
L3/CR1	0.32	-
SINEs	13.64	10.74
B1	10.74	10.74
B2	-	-
B4/RSINE	-	-
ID	-	-
MIR/MIR3	2.90	-
Elementos LTR	8.55	4.09
ERV claseI	2.92	2.02
ERV claseII	0.30	0.30
ERV claseIII	1.55	0.19
MaLRs (III)	3.78	1.58
Elementos de DNA	3.03	1.00
Charlie	1.41	0.14
Otros hATs	0.31	-
Tigger	1.06	0.76
Mariner	0.10	0.07
Sin clasificar	0.15	0.14
Total	46.36	24.05

Tabla 4: Proporción de elementos repetidos en el genoma humano. Tipos de secuencias repetidas y su porcentaje total en el genoma humano (adaptado de Lander *et al.*, 2001). hATs: ; MIR: ; MaLRs: .

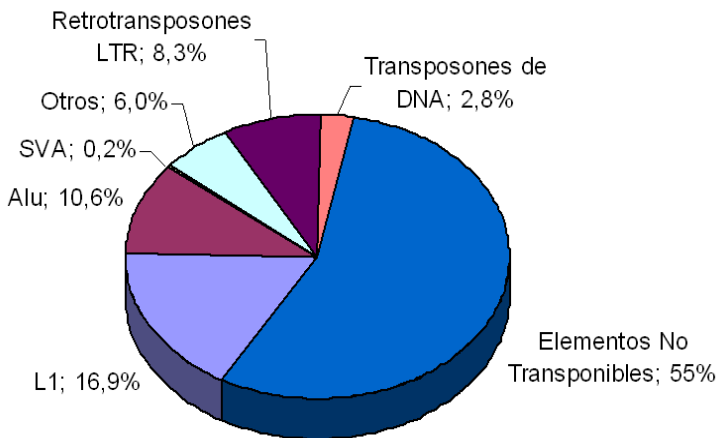


Figura 7: Elementos repetidos en el genoma humano. Un 45% del genoma humano está compuesto por secuencias repetidas.

Casi un 3% del genoma humano está compuesto por transposones de DNA de cortar y pegar, concretamente encontramos elementos Charlie, Mariner y Tigger (Lander *et al.*, 2001; Goodier & Kazazian, 2008; Mills *et al.*, 2007), aunque ya no son activos. Los motivos por los que estos elementos han perdido su capacidad de movilizarse en el genoma humano no se conocen por completo, pero las evidencias genéticas indican que dejaron de ser móviles en nuestro genoma hace millones de años.

Dentro de los retrotransposones encontramos las dos clases principales: los de tipo LTR, representados mayoritariamente por los denominados hERVs (por *human endogenous retroviruses*), que parecen no ser activos, y los de tipo no-LTR, que constituyen por sí solos un tercio de nuestro genoma (Lander *et al.*, 2001). Precisamente

dentro de este segundo tipo encontramos un grupo de elementos que continúa movilizándose de forma activa dentro de nuestro genoma, lo que los hace especialmente relevantes para su estudio en biomedicina, dado que su movilización puede provocar la aparición de enfermedades genéticas tales como por ejemplo la hemofilia (Kazazian *et al.*, 1988) o la distrofia muscular de Duchenne (Narita *et al.*, 1993).

LINE-1 es el único retrotransposón autónomo activo en la actualidad en el genoma humano, y sus inserciones *de novo* continúan impactando en nuestro genoma. Además, su actividad es responsable de la movilización del resto de elementos activos no autónomos, constituidos por Alu, SVA y los pseudogenes procesados o retrogenes (figura 8) (revisado en Goodier & Kazazian, 2008).

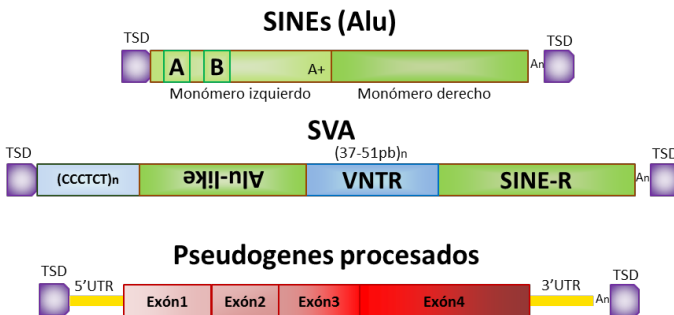


Figura 8: Retrotransposones no autónomos en el genoma humano movilizados por LINE-1. Esquema de las secuencias movilizadas por L1: **SINEs** (*Short INterspersed Elements*), **Alu**: A, caja A; B, caja B; A+, región rica en adeninas. **SVA** (SINE-VNTR-Alu): Alu-like, región similar a Alu; VNTR (por *Variable Number Tandem Repeats*), repeticiones en tándem de número variable. (Adaptado de Goodier & Kazazian, 2008 y Mills *et al.*, 2007)

1.2.- Los elementos LINE-1 (L1) en el genoma humano.

1.2.1.- Descubrimiento.

A finales de los años 60 ya se había descrito la existencia de secuencias repetidas en los genomas de eucariotas, estimándose que representaban entre un 20-60% del tamaño total del genoma. Este tipo de secuencias se dividían entre secuencias altamente repetidas y secuencias moderadamente repetidas (Britten & Kohne, 1968); unos años después, se encontró por primera vez una secuencia repetida inusualmente larga (de aproximadamente 6.4kb) en las proximidades del gen de la beta-globina (Adams *et al.*, 1980).

Años más tarde se describió que esta secuencia estaba compuesta por cuatro subfamilias distintas, que pertenecían a una misma familia de secuencias repetidas en el genoma, denominadas en ese momento como *Kpn I* (Shafit-Zagardo, 1982); concretamente pertenecían a las subfamilias *Kpn I* 1.2kb, 1.5kb y dos subfamilias de *Kpn I* 1.8kb. Se comprobó además que estas secuencias aparecían en el mismo orden (5'-pBK(1.8)26-pBK(1.5)54-pBK(1.2)11-pBK(1.8)11-3'), conformando una secuencia repetida con un tamaño total de 6.4kb (Shafit-Zagardo, 1982; Thayer & Singer, 1983). Además, se comprobó que esta secuencia aparecía dispersa por el genoma, presentaba múltiples polimorfismos y, al menos en algunos casos, podía

transcribirse a RNA (DiGiovanni *et al.*, 1983; Lerman *et al.*, 1983).

Esta secuencia se caracterizó con más detalle y pasó a conocerse como LINE-1 (Singer, 1982; Grimaldi, 1984). Poco después, en el citoplasma de varias líneas celulares, se descubrió la existencia de transcritos de estas secuencias de hasta 6,5kb de longitud (Skorowski & Singer, 1985), demostrándose así que era capaz de expresarse.

Cómo estas secuencias se habían hecho tan abundantes en el genoma seguía siendo un misterio, pero un análisis más exhaustivo reveló otra gran sorpresa: la existencia de una región dentro de los elementos LINE-1 que codificaba una proteína similar a otras reversotranscriptasas conocidas (Muranyi, 1988); pruebas adicionales apuntaban que ésta podía seguir siendo funcional (Sakaki, 1987).

Así, el descubrimiento de que en humanos este tipo de elemento continuaba movilizándose de forma activa se produjo poco después, y coincidiendo con la descripción del primer caso de una enfermedad humana causada por LINE-1, la hemofilia A. Se detectó la presencia de dos inserciones distintas (de 3.8 y 2.3kb de longitud) de este elemento en el exón 14 del gen del factor VIII, en 2 pacientes de hemofilia A (de un total de 240) (Kazazian, 1988). Años más tarde, se consiguió aislar el elemento completo de estos individuos, y se comprobó que se trataba de dos elementos del mismo tamaño pero parcialmente distintos a nivel de secuencia, denominándose L1.1 y L1.2. En el mismo estudio se consiguió además aislar

un tercer elemento también distinto, llamado L1.3 (Dombroski *et al.*, 1991).

Años más tarde, en el laboratorio del Dr. Haig Kazazian, el también Dr. John V. Moran, aprovechando las propiedades del retrotransposón y mediante el empleo de un ensayo de gen reportero, diseñó un ensayo pionero para poder estudiar en células en cultivo la capacidad de movilizarse de estos elementos, (Moran *et al.*, 1996); el Dr. Moran utilizó un casete diseñado originalmente para el estudio de la movilidad de retrovirus (Heidmann *et al.*, 1988; Freeman *et al.*, 1994) y lo adaptó al elemento LINE-1. Sorprendentemente, este sistema no sólo demostró que el elemento es activo, sino que además su tasa de movilización es muy alta en líneas celulares humanas transformadas (Moran *et al.*, 1996). Desde entonces, este sistema de movilidad *in vitro* se ha utilizado por numerosos laboratorios del mundo para descifrar procesos de la biología de LINE-1.

1.2.2.- Estructura.

Los elementos LINE-1 están compuestos por una región 5' no traducida (5' UTR), que contiene actividad promotora interna (Swergold, 1990), y una actividad promotora antisentido (ASP) menos potente (Speck, 2001) (figura 9). Seguidamente se encuentra un marco abierto de lectura bicistrónico que codifica las proteínas responsables de la movilidad de L1, conocidas como ORF1 y ORF2 (Mcmillan & Singer, 1993), descritas en detalle posteriormente; entre ambas proteínas existe una

pequeña región intergénica que puede ser modificada sin afectar en gran medida la retrotransposición (Alisch *et al.*, 2006). Tanto ORF1 como ORF2 son requeridas para la movilización del elemento (Feng *et al.*, 1996; Kulpa & Moran, 2005). Finalmente, contienen una región 3' no traducida (3' UTR por *Untranslated Region*), que termina en una señal de poliadenilación relativamente débil, lo cual hace que sea frecuente encontrar transducciones de otras secuencias situadas aguas abajo del elemento tras su movilización (Moran *et al.*, 1999; Pickeral *et al.*, 2000). Cuando se analiza el genoma, lo más frecuente es encontrar estos elementos flanqueados por duplicaciones del sitio diana de unas 2 a 20 pares de bases (Kazazian *et al.*, 1988).



Figura 9: Estructura de un elemento LINE-1 humano. Representación esquemática de las secuencias contenidas en el retroelemento: UTR, región sin traducir; ORF, marco abierto de lectura; An, cola de poliAdeninas.

En la región 5'UTR, además de las actividades promotoras, encontramos sitios de unión de diversos factores de transcripción, de entre los que destacan YY-1 (Athanihar *et al.*, 2004) y RUNX3 (Yang *et al.*, 2003). La región 5'UTR de L1 contiene también múltiples sitios susceptibles de sufrir metilación (islas CpG), especialmente importantes para la célula para ejercer un control de la expresión del elemento. La región 5'UTR se encuentra fuertemente metilada en células diferenciadas, mientras que el nivel de

metilación es bajo en células indiferenciadas, como ocurre por ejemplo al reprogramar las células a células pluripotentes inducidas (o iPSCs) (Wissing *et al.*, 2012).

ORF1 (figura 10) es una proteína de aproximadamente 40kDa con actividad chaperona de ácidos nucleicos (Martin & Buschman, 2001). En la región N-terminal, necesaria para la interacción de la proteína con otras proteínas, destaca la presencia de un dominio de cremallera de leucina (*Leucine Zipper* o LZ) (Holmes *et al.*, 1992; Martin, 2003). A continuación se encuentra un motivo de unión a RNA denominado RRM (*RNA Recognition Motif*) (Khazina *et al.*, 2009), seguido de un dominio C-terminal (CTD) y la cooperación de ambos se requiere para su unión al RNA (Khazina *et al.*, 2009).

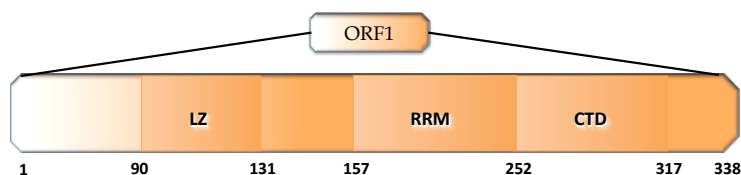


Figura 10: Estructura de ORF1 de un elemento LINE-1 humano. Los números indican las posiciones de aminoácidos de inicio y final de cada dominio. LZ, cremallera de leucina; RRM, dominio de reconocimiento de RNA; CTD, dominio C-terminal.

En la traducción del mensajero de LINE-1 se producen múltiples copias de ORF1, que forman homotrímeros mediante la interacción cola con cola (*coiled-coil* o C-C) de su parte N-terminal, estos

trímeros se unen al RNA mediante sus dominios RRM y CTD, formando así parte de la RNP completa de LINE-1 (Martin, 2003; Khazina *et al.*, 2011, revisado en Martin, 2006).

ORF2 (figura 11) es una proteína multidominio de aproximadamente 150kDA. Está compuesta por los dominios endonucleasa (EN), un dominio intermedio llamado dominio Z de función desconocida (Clements & Singer, 1998), un dominio reversotranscriptasa (RT) y un motivo C-terminal rico en cisteínas (C) (Clements & Singer, 1998).

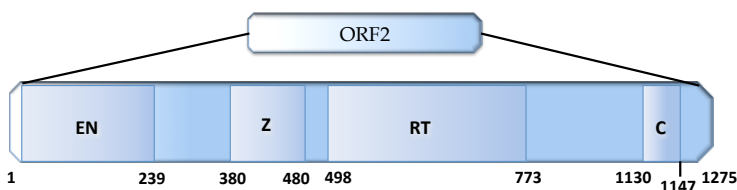


Figura 11: Estructura de ORF2 de un elemento LINE-1 humano. Los números indican las posiciones de aminoácidos de inicio y final de cada dominio; EN, dominio endonucleasa; Z, dominio Z; RT, dominio reversotranscriptasa; C, dominio rico en cisteínas. Adaptado de Clements & Singer, 1998.

A destacar, mutaciones en los tres dominios principales identificados en L1-ORF2p (EN, RT y C), indican que son actividades necesarias para la retrotransposición (Moran *et al.*, 1996), aunque se desconocen bastantes detalles de su mecanismo de acción.

1.2.3.- Mecanismo de movilización de L1: TPRT (*Target Primed Reverse Transcription*).

Los retrotransposones L1 son capaces de hacer una copia de sí mismos en otro lugar del genoma, permaneciendo en su localización original en el DNA. Para ello, a través de un RNA intermediario, son capaces de retrotranscribir su secuencia a una nueva localización (Singer, 1982).

Los elementos L1 durante su moviliación producen un corte en una de las cadenas del DNA (denominada cadena inferior o *bottom-strand*) (Feng *et al.*, 1996), y utilizando como cebador (o *primer*) el extremo 3'-OH que queda liberado tras el corte y como molde el RNA del elemento comienza la reversotranscripción, luego se sintetiza la cadena complementaria y una ligasa celular sella los huecos (Luan *et al.*, 1993). Las fases finales de la integración aun no están muy claras, pero es frecuente encontrar duplicaciones cortas del sitio de inserción (*Target Site Duplications* o TSDs) de unas 2 a 20 pares de bases flanqueando a estos elementos, siendo estas TSDs muy características de este mecanismo de inserción (Jurka *et al.*, 1997). El proceso se resume en 4 etapas cuyo esquema se muestra en la figura 12 (Luan *et al.*, 1993) y que a continuación de detallan:

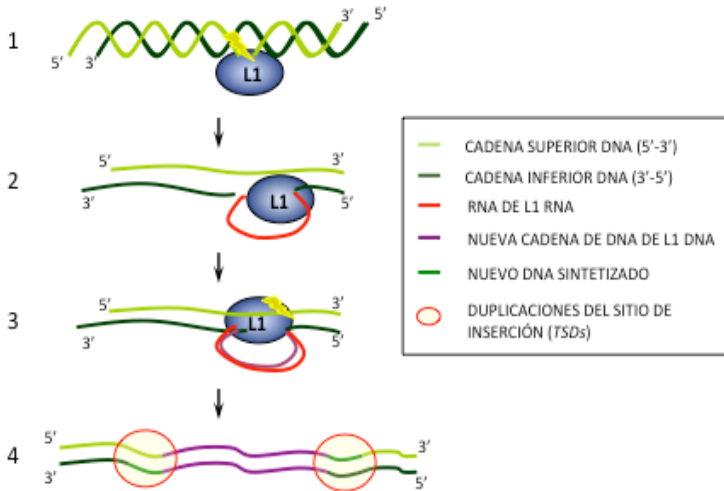


Figura 12: Target Primed Reverse Transcription (TPRT). Mecanismo de integración de L1 en el genoma (detalles en el texto). Adaptado de Luan *et al.*, 1993.

1. La endonucleasa del L1 corta en una cadena del DNA (cadena inferior o *bottom strand*) donde reconoce la secuencia consenso 5' TTTT/A 3', o una secuencia similar, revelando el extremo 3'-OH de la base.
2. Usando como molde el RNA mensajero de L1, y como cebador el extremo 3'-OH libre del sitio donde se dará la inserción, se inicia la síntesis de la primera cadena de DNA del retroelemento, empezando por la cola de poliA situada en el extremo 3' del RNA.
3. Mientras tiene lugar la síntesis de cDNA de L1, se produce un corte en la segunda cadena (cadena superior o *top strand*), pero normalmente no se produce en la misma posición que el primer corte.

4. Se va sintetizando la cadena complementaria de DNA de L1, a la vez que se sintetiza el DNA del sitio de inserción; el corte de la cadena superior se produce aguas abajo con respecto a la posición del primer corte (cadena inferior, 3'-5').

5. Al actuar una polimerasa a ambos lados de la inserción, se duplica la información parcial, dando lugar a duplicaciones del sitio de inserción o TSDs (*Target Site Duplications*), generalmente de 2 a 20 pares de bases de longitud, características del mecanismo de inserción por TPRT.

Es importante destacar que, por las propiedades de la reversotranscriptasa del L1, existe la posibilidad de que sea el propio L1 el que, mediante un cambio de molde (*template switching*), pase a utilizar como cebador el extremo que queda libre de la cadena superior para continuar sintetizando DNA en la cadena inferior, que posteriormente sería sellada por una ligasa. Sin embargo, a día de hoy no se conoce si es una polimerasa celular o la propia RT del L1 la encargada de realizar las síntesis de las cadenas complementarias.

Asimismo, las propiedades de las ribonucleopartículas de L1, como el "*template switching*", permiten la aparición de inserciones que contienen dos fragmentos truncados de L1 enfrentados en direcciones opuestas, con una inversión de parte de la secuencia, dichos eventos proceden de una reacción de "doble síntesis" o "*twin priming*" (Ostertag *et al.*, 2001).

1.2.4.- Consecuencias de la actividad de L1.

En el genoma humano, las secuencias repetidas móviles activas o inactivas suponen aproximadamente un 45% del total (Cordaux & Batzer, 2009); y aunque la mayoría no tienen ya capacidad para moverse, otras sí que conservan el potencial de movilizarse (revisado en Carreira *et al.*, 2014), pudiendo tener numerosos efectos. No obstante, para asegurar su continuidad evolutiva, las secuencias repetidas deben movilizarse en las células germinales y/o durante el desarrollo embrionario temprano. Cada día hay más evidencias de que la movilidad en las células germinales es baja o casi inexistente, de manera similar a lo que ocurre en muchas células somáticas adultas (Belancio *et al.*, 2010), mientras que en las células embrionarias, donde el genoma está demetilado a nivel general, los niveles de expresión y movilidad son mucho más altos (García-Pérez *et al.*, 2007). De hecho, estudios posteriores usando un modelo murino de movilidad de un elemento LINE-1 humano revelaron que la mayoría de inserciones heredables se producen durante el desarrollo embrionario, mientras que las inserciones en células germinales son raras (Kano *et al.*, 2009).

No obstante, el hecho de que estas secuencias se movilizan durante el desarrollo embrionario se traduce, en la mayoría de las nuevas inserciones, en un mosaicismo en el individuo (García-Pérez *et al.*, 2007). Las nuevas copias de secuencias repetidas se encontrarán sólo en algunas células y sus descendientes y tan sólo algunas de estas inserciones pasarán a la siguiente generación. De hecho, mediante

la caracterización de nuevas inserciones que producen una enfermedad genética humana, se ha demostrado que éstas se acumulan durante estadios tempranos del desarrollo y que los seres humanos somos un mosaico para nuestro contenido en LINE-1 (Van der Huk *et al.*, 2007). Por otro lado, durante la vida de un individuo siguen existiendo células madre o progenitoras que podrían albergar nuevos eventos de retrotransposición, tal y como se ha descrito en células progenitoras neuronales (Muotri *et al.*, 2005), aunque se desconoce que otras células madre adultas podrían albergar movilidad de LINE-1.

Dado que el sitio de inserción del elemento es casi aleatorio, LINE-1 se puede insertar en prácticamente cualquier localización del genoma. Esto quiere decir que en ocasiones su inserción puede producirse en un intrón o en un exón afectando a la expresión y/o a la función de los genes situados en el área cercana (Beck *et al.*, 2011).

Asimismo, en el elemento LINE-1 existen distintas secuencias reguladoras, tales como sitios de unión de factores de transcripción, una señal de poliadenilación y dos actividades promotoras distintas (sentido y antisentido). Esto provoca que una nueva inserción de LINE-1 pueda afectar en gran medida al funcionamiento de las zonas donde se integra, alterando la expresión de genes situados en zonas próximas de muchas maneras diferentes (revisado en Beck *et al.*, 2011).

A los efectos intrínsecos a la secuencia nucleotídica de LINE-1 además hay que añadirle el efecto de su movilidad en células indiferenciadas, donde el sitio de inserción es silenciado epigenéticamente mediante deacetilación de histonas

(García-Pérez *et al.*, 2010). Este fenómeno es específico para el mecanismo de inserción de los elementos LINE (TPRT) y no ocurre en el caso de inserciones virales (García-Pérez *et al.*, 2010); por tanto la actividad de LINE-1 también es capaz de alterar la epigenética la región donde se retrotranscribe.

Todo esto supone que la importancia de estos elementos no queda restringida al plano evolutivo, sino también a nivel biomédico, ya que su actividad puede afectar a la aparición o al desarrollo de enfermedades genéticas. Actualmente se conocen 65 enfermedades en humanos causadas por las inserciones de L1, Alu y SVAs, y aunque dichas enfermedades son causadas por lo que podríamos considerar una mutación puntual, se ha descrito que la actividad de LINE-1, ya sea por su movilización o la movilización de Alu y SVA, puede contribuir al desarrollo y progresión de tumores (Iskow *et al.*, 2010).

Otra de las consecuencias ocasionales de la movilidad de LINE-1 es la pérdida de información genética, pudiendo causar deleciones en el sitio de inserción (Gilbert *et al.*, 2002). Estas deleciones pueden ser de tamaño considerable, como en el caso de una inserción de L1 en el componente X del gen de la piruvato deshidrogenasa (PDHX) donde el tamaño de la deleción producida es de 46 kb, o por poner otro ejemplo producir la deleción de un gen entero como el HLA-A tras una inserción de SVA (Gilbert *et al.*, 2002; Takasu, 2007).

Además, es relativamente frecuente (entre un 10% y un 20% de las ocasiones), que durante la

expresión del retroelemento LINE-1 la señal de poliadenilación, que es relativamente débil, no sea capaz de frenar la transcripción (Moran *et al.*, 1999). A el mensajero formado se le añade un fragmento de DNA genómico, independiente del elemento, que puede ser movilizado igualmente en la actividad del retrotransposón. Este fenómeno es conocido como transducción en 3' (Moran *et al.*, 1999).

También, aunque de manera menos frecuente, se puede dar la transcripción del elemento, ya sea LINE-1, Alu o SVA, desde un promotor situado aguas arriba, dando lugar a transducciones en 5' (Damert *et al.*, 2009).

Si sumamos estas dos actividades podemos considerar que la actividad de LINE-1 puede servir para producir un “barajamiento” de exones (Moran *et al.*, 1999), pudiendo dar lugar a la aparición de nuevos genes, o bien nuevas características en genes preexistentes (Esnault *et al.*, 2000). Un claro ejemplo de este fenómeno lo encontramos en la transducción de la secuencia entera del gen AMAC1 (*acyl-malonyl condensing enzyme 1*) en la movilización de elementos SVA, que ha generado múltiples copias transcripcionalmente activas del gen en el genoma humano (Xing, 2006).

Otra manera que han tenido los retrotransposones de afectar los genomas en los que habitan es la recombinación. Debido a la enorme cantidad de estas secuencias presentes en el genoma, es frecuente que se produzca recombinación homóloga entre dos elementos del mismo tipo situados en dos regiones distintas del genoma (Jinks-Robertson & Petes, 1986). Esta recombinación es especialmente frecuente en el caso de secuencias Alu,

donde podemos encontrar hasta 50 ejemplos de recombinación de estos elementos que causan la aparición de una enfermedad genética, como por ejemplo hipercolesterolemia (Lehrman *et al.*, 1987), la distrofia muscular de Duchenne (Hu *et al.*, 1991) o leucemia de linaje mixto (MLL, Aplan, 2006).

Además, en el genoma humano encontramos un enriquecimiento importante de elementos Alu en las cercanías de duplicaciones segmentales, algo que en el caso de ratones ocurre para elementos LINE1 y retrotransposones LTR, pero no con SINEs (Bailey *et al.*, 2006; She *et al.*, 2008).

No obstante, la actividad de L1 también puede ser causante de efectos beneficiosos para el individuo; el caso mejor conocido es el de la fusión mediada por la retrotransposición del gen de la ciclofilina A en el gen TRIM-5, formando un nuevo gen llamado TRIMcyp, que aporta resistencia ante el virus HIV en *Aotus* (Sayah *et al.*, 2004).

1.2.5.- Regulación de LINE-1.

Dado el inherente potencial mutagénico de la actividad de estos elementos, el control de su movilidad resulta esencial para mantener la estabilidad del genoma. Dicho control se efectúa a varios niveles, empezando por el control de su transcripción. Para ello, principalmente en células somáticas, se produce una fuerte metilación de la región 5'UTR del elemento. De esta manera el promotor queda silenciado, y por lo tanto sus niveles de expresión suelen ser muy bajos (Hata & Sakaki, 1997).

No obstante, en células indiferenciadas, tales como células progenitoras y células embrionarias, los niveles de expresión de LINE-1 suelen ser altos, y generalmente la metilación del promotor no se produce o es débil (García-Pérez *et al.*, 2007), como ocurre por ejemplo durante la embriogénesis, donde se produce una demetilación generalizada del genoma (Fadloun *et al.*, 2014). Ante estos casos, en los que la transcripción del elemento no puede silenciarse mediante la metilación de su promotor, el genoma humano dispone de una segunda línea de defensa, que consiste en el uso de micro RNAs (Aravin *et al.*, 2007; Heras *et al.*, 2013) para evitar la expresión del elemento (figura 13).

No obstante, algunas células progenitoras a pesar de presentar niveles de expresión muy elevados, no expresan las proteínas necesarias para desarrollar el mecanismo de silenciamiento previamente descrito; es el caso, por ejemplo, de las células progenitoras neuronales o NPCs (por *Neural Progenitor Cells*), en las que parece que el control de la expresión de LINE-1 se ejerce mediante el represor transcripcional MeCP2 (por *Methyl Cytosine binding Protein 2*) (Muotri *et al.*, 2010).

Asimismo otras proteínas, como MOV10L1, también parecen limitar la movilidad del elemento. En concreto, esta proteína actúa en la espermatogénesis, colaborando con la maquinaria de formación de piRNAs para prevenir la movilización de los retrotransposones (L1, Alu, SVA). La ausencia de MOV10L1 provoca un aumento notable de la expresión de L1 y un bloqueo de la espermatogénesis (Frost *et al.*, 2010), aunque no se han detectado

niveles de expresión significativos de esta proteína en placenta o testículos (Reichman *et al.*, 2013).

MOV10 es una helicasa de RNA que forma parte del complejo silenciador RISC (por *RNA Induced Silencing Complex*). Se ha encontrado colocalizada con la ribonucleopartícula de L1 en gránulos de estrés en distintos tipos celulares y se ha demostrado que inhibe la expresión y retrotransposición del elemento, aunque aún no está claro si esta inhibición se produce a nivel transcripcional o post-transcripcional (Goodier *et al.*, 2013).

Tex 19.1 es otra de las proteínas que parecen estar implicadas en el control de la movilidad del retroelemento, aunque curiosamente, esta proteína se expresa en la placenta, y no en el embrión (Reichman *et al.*, 2013); su deficiencia provoca un notable aumento de la expresión de LINE-1 en la placenta y conlleva retrasos en el desarrollo embrionario intrauterino en ratones (Reichman *et al.*, 2013).

Aun así, si el elemento es capaz de formar una ribonucleopartícula funcional, existe una tercera vía para controlar su movilidad, se trata de las proteínas APOBEC, conocidas por su actividad antiviral, especialmente frente al virus HIV. Concretamente su mecanismo de acción consiste en editar el RNA del virus mediante deaminación (Mangeat *et al.*, 2003). Estas proteínas también son capaces de controlar la movilidad de los retroelementos (Bogerd *et al.*, 2006; Wissing *et al.*, 2012), si bien en este caso su mecanismo de inhibición parece funcionar incluso sin la participación de su actividad deaminasa (Horn *et al.*, 2013).

Además de los mecanismos de defensa ya mencionados, en células indiferenciadas existe otro

nivel de control adicional, en el caso de que el elemento sea capaz de completar la retrotransposición. Este mecanismo consiste en silenciar las secuencias introducidas por el retroelemento mediante deacetilación de las histonas de la región, produciéndose un tipo de marca en el DNA que puede mantenerse incluso tras la diferenciación (García-Pérez *et al.*, 2010).

Este proceso se puede revertir mediante la adición de inhibidores de las deacetilasas de histonas (*Histone deacetylases* o HDACs). Este mecanismo de silenciamiento se produce específicamente ante secuencias introducidas mediante la TPRT, y no se produce en el caso de las inserciones retrovirales (García-Pérez *et al.*, 2010). Es de destacar que, en el caso de las células indiferenciadas, si se retira el inhibidor de las HDACs, se vuelve a producir el silenciamiento de la región; sin embargo, esto no ocurre después de que las células se hayan diferenciado, es decir, en este caso al retirar el inhibidor la región no se vuelve a silenciar (García-Pérez *et al.*, 2010). Los mecanismos detrás de esta especificidad de marca epigenética aún no se han descubierto, pero este fenómeno pone de manifiesto la importancia del elemento a la hora de afectar la expresión de las zonas donde se integra en células indiferenciadas.

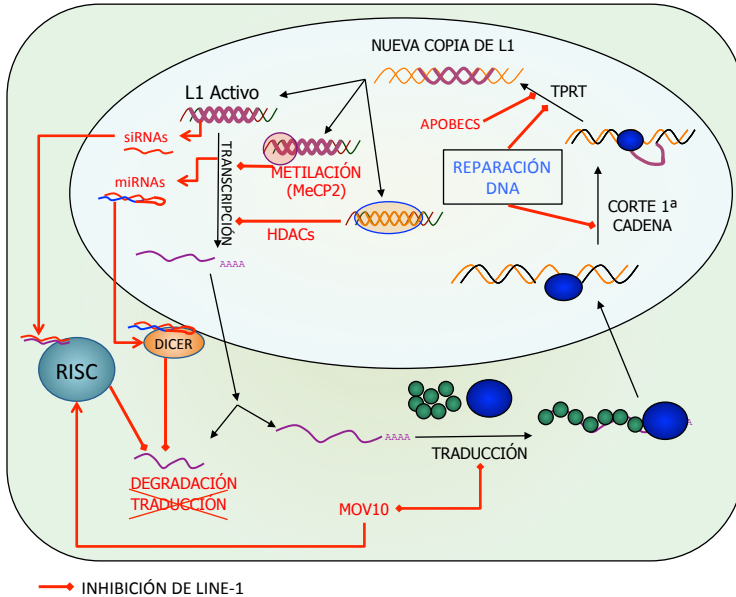


Figura 13: Vías de control de la movilidad de LINE-1. Represión de la transcripción mediante metilación de DNA, mediada por la proteína MeCP2, o por deacetilación de las histonas de la región mediante HDACs. Degradación del RNA mediante siRNAs y miRNAs. Inhibición de la formación de ribonucleopartículas mediante MOV10. Reparación de las roturas causadas por la endonucleasa del elemento mediante la maquinaria de reparación del DNA. Inhibición de la TPRT mediante las proteínas APOBECs.

1.2.6.- Vía de retrotransposición de LINE-1 independiente de la endonucleasa (ENi) y reparación del DNA.

En el año 1996, además de realizarse por primera vez el ensayo clásico de retrotransposición (Moran *et al.*, 1996), se demostró que un dominio endonucleasa funcional era necesario para que se produjera la retrotransposición (Feng *et al.*, 1996), es decir, que el elemento L1 sólo es capaz de moverse mediante el mecanismo clásico de TPRT descrito anteriormente (Luan *et al.*, 1993).

Sin embargo, unos años más tarde, en el laboratorio del Dr. Moran, se describió que el retroelemento es capaz de moverse, incluso sin un dominio endonucleasa funcional, en células CHO (por *Chinese Hamster Ovary*) deficientes en la reparación del DNA por la vía de unión de extremos no homólogos (*Non-Homologous End Joining* o NHEJ) (Morrish *et al.*, 2002).

Posteriormente, se describió que además el retroelemento es capaz de insertarse en regiones teloméricas y subteloméricas, en células cuyos telómeros se encuentran desprotegidos por una deficiencia en las proteínas DNA-PKs de la ruta NHEJ, incluso teniendo mutado su dominio endonucleasa (Morrish & García-Pérez *et al.*, 2007).

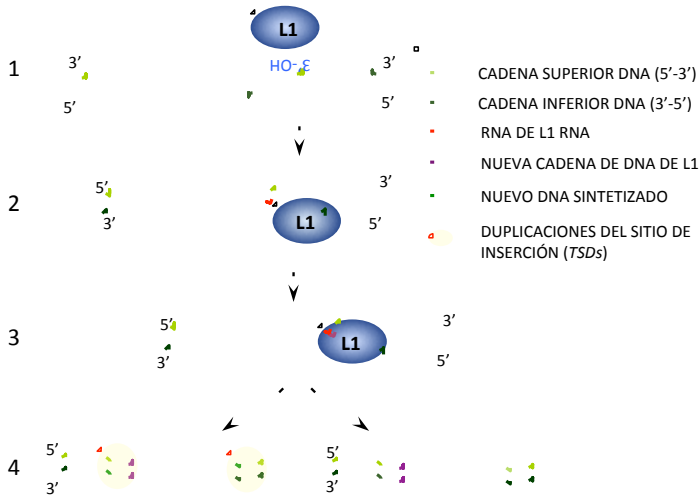


Figura 14. Retrotransposición independiente de endonucleasa (ENi o *Endonuclease independent retrotransposition*). El proceso de retrotransposición transcurre de manera casi idéntica al mecanismo de TPRT, con la diferencia de que el extremo 3' hidroxilo libre necesario para iniciar el proceso no es generado por la actividad endonucleasa del L1.

En estudios en el genoma humano, se describió que los elementos L1 presentan un alto grado de variabilidad en su secuencia. Uno de estos estudios determinó que ciertas mutaciones puntuales en cualquiera de los 3 dominios fundamentales (EN, RT y C) podían anular la movilidad del elemento, pero describió también que otras, a pesar de afectar a los niveles relativos de retrotransposición, no impiden su movilización (Seleme *et al.*, 2006). Asimismo, en un estudio posterior, Sen y colaboradores analizaron inserciones en el genoma humano que presentaban características anómalas, denominándolas como

“NCLI” por “*Non Classical L1 insertions*”, y estimaron que el número total de inserciones atípicas presentes en el genoma humano podía ascender hasta 2000-2800 desde que se produjo la radiación de los primates (Sen *et al.*, 2007). Estos autores describieron que estas inserciones atípicas se producían mediante regiones de microhomología entre el RNA del elemento y roturas en el genoma, actuando como sustrato para reparar roturas de doble cadena en el DNA, que se produce sin intervención de la endonucleasa del elemento, y concluyeron que esta vía de retrotransposición ha podido tener un alto impacto a la hora de dar forma al genoma humano actual, así como que existe una relación entre los retroelementos y la reparación de roturas de doble cadena del DNA (Sen *et al.*, 2007).

Años más tarde, Beck describió algunos ejemplos de movilidad por esta vía en el genoma humano (Beck *et al.*, 2010), demostrando definitivamente que la vía de retrotransposición independiente de endonucleasa, o vía de retrotransposición ENi, es una vía alternativa de movilización de L1 que continua produciéndose de forma activa en nuestros genomas (Beck *et al.*, 2010).

Otra gran sorpresa se produjo al analizar las inserciones que se producían como consecuencia de la vía de retrotransposición ENi, ya que las características de éstas eran muy diferentes, en muchos casos, a las que se observan en inserciones producidas cuando el retroelemento se moviliza por la vía clásica (Moran *et al.*, 1996; Gilbert *et al.*, 2002), dependiente de su endonucleasa; así, mientras que normalmente se aprecia la presencia de una cola de poliA en el extremo 3' de una inserción típica, en el

caso de las inserciones por la vía ENi dicha cola no siempre está presente, y de hecho es frecuente encontrar elementos truncados en 3' (Morrish *et al.*, 2002; Gilbert *et al.*, 2002; Symer *et al.*, 2002). También resulta llamativa la ausencia de las duplicaciones que suelen aparecer flanqueando el elemento (TSDs), e incluso ocasionalmente se observan deleciones en el sitio de inserción (Gilbert *et al.*, 2002). Por último, es frecuente encontrar estas inserciones acompañadas de secuencias cortas de DNA que parecen proceder de otros cDNAs endógenos (Gilbert *et al.*, 2005; Morrish *et al.*, 2002).

Por tanto, dada su capacidad de causar alteraciones cromosómicas en la región donde se inserta, controlar la actividad de L1 es de vital importancia para el correcto desarrollo de la célula y el individuo. Estas alteraciones pueden incluir deleciones (Gilbert *et al.*, 2002), reordenaciones de material genético (Moran *et al.*, 1999), o promover traslocaciones al recombinarse dos secuencias homólogas, principalmente de L1 o Alu, como se describió anteriormente. Por tanto, la actividad de LINE-1, que incluso en condiciones normales es una fuente potencial de inestabilidad genómica (Gilbert *et al.*, 2002; Symer *et al.*, 2002), puede aumentar su impacto en el genoma ante deficiencias en los mecanismos de reparación (Morrish *et al.*, 2002, Morrish & García-Pérez *et al.*, 2007).

Ahora bien, sabiendo que la movilidad de L1 por la vía ENi se produce de forma natural, y que su frecuencia aumenta enormemente ante fallos en la reparación del DNA por la vía de NHEJ, nos

planteamos qué ocurriría en el caso de otras deficiencias en la reparación del DNA, en las que los niveles de movilidad del retroelemento y la vía de retrotransposición ENi pueden estar igualmente desreguladas.

Entre multitud de posibles modelos elegimos para la realización de esta tesis la ruta de reparación de la Anemia de Fanconi, por motivos que expondré a lo largo de los siguientes apartados.

1.3.- Anemia de Fanconi (*Fanconi anemia* o FA).

1.3.1.- Introducción

La Anemia de Fanconi fue descrita por primera vez por el Dr. Guido Fanconi en el año 1927. Describió el caso de una familia con 3 miembros que presentaban distintos tipos de anomalías físicas, y que finalmente fallecieron a causa de un tipo de anemia perniciosa (Fanconi, 1927). Existen muchos genes cuyas mutaciones provocan la aparición de la enfermedad, descritos como grupos de complementación que se han ido descubriendo con el paso de los años (revisado en Kee & D'andrea, 2012). Actualmente el número de genes cuya mutación bialélica causa la enfermedad asciende a 15. Estos genes constituyen la denominada ruta de la anemia de Fanconi (*FA pathway* o *FA/BRCA pathway*) y se conocen como: FANCA, FANCB, FANCC, FANCD1 (BRCA2), FANCD2, FANCE, FANCF, FANCG, FANCI, FANCI (BRIP1), FANCL, FANCM, FANCN (PALB2), FANCO (RAD51C) y FANCP (SLX4).

La anemia de Fanconi es una enfermedad genética recesiva, caracterizada por ser un síndrome de inestabilidad cromosómica, cuya frecuencia se estima en 1 de cada 100.000 nacimientos (Rosenberg *et al.*, 2011), y aproximadamente un 10% de los pacientes son mosaico para la enfermedad (Hoehn, 2002), con una edad media de supervivencia de 20-25 años, aunque la edad varía en función de las

distintas complicaciones que aparecen con la enfermedad (revisado en Alter, 2003).

La incidencia de portadores de un alelo causante de la enfermedad se estimaba en torno a 1 de cada 300 nacimientos (Swift *et al.*, 1971), aunque estudios más recientes demuestran que esta proporción puede variar notablemente entre poblaciones (Kutler *et al.*, 2004), y se sitúa entre 1:156 – 1:209 en la población americana y 1:66 – 1:128 en la población israelí (Rosenberg *et al.*, 2011), lo que sugiere que la proporción general de portadores de un alelo mutante puede ser mucho más elevada de lo que se consideraba.

La principal causa de la aparición de la sintomatología propia de la enfermedad es la deficiencia en la coordinación de las diferentes rutas de reparación del DNA, lo que conlleva un aumento en la acumulación de daños de diversa índole, derivados del propio metabolismo de la célula, o por factores ambientales como el estrés oxidativo o los diferentes tipos de radiación (figura 15).

Toda esta acumulación de daños en el DNA se traduce en una elevada inestabilidad genómica en las células de estos pacientes, que suele manifestarse en forma de anomalías físicas, fallo progresivo de la médula ósea y aparición de cáncer y/o tumores sólidos (Figura 15). Destaca una especial sensibilidad a los daños causados por agentes intercalantes de DNA, tales como el cisplatino, la mitomicina C, el diepoxibutano (DEB) o el melphalan, característica que se emplea para su diagnóstico y en su estudio en el laboratorio (Carreau *et al.*, 1999).

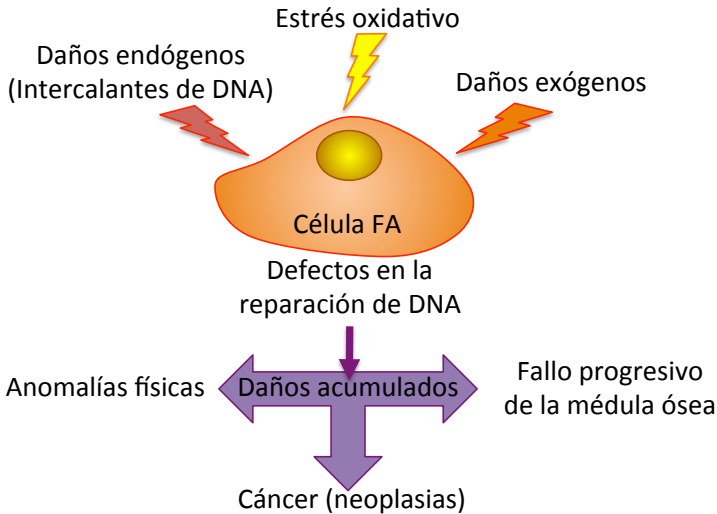


Figura 15. Causas y consecuencias de las deficiencias en las células Fanconi. Los daños producidos en el DNA por el funcionamiento normal de la célula, así como los que surgen por factores ambientales, no se reparan correctamente en las células de Anemia de Fanconi. Se produce una acumulación de mutaciones, que causa anomalías físicas en el momento del nacimiento, pérdida progresiva de la función de la médula ósea, y en última instancia desarrollo de leucemias y/o tumores sólidos (Adaptado de Kee & D'andrea, 2012).

1.3.2.- Sintomatología.

Una de las principales características de la sintomatología propia de la anemia de Fanconi es que es muy heterogénea; en la mayoría de las ocasiones, las manifestaciones de la enfermedad incluyen anomalías físicas (60-75% de los pacientes), elevada predisposición a desarrollar distintos tipos de cáncer o tumores (aproximadamente 24% de los pacientes) y un fallo progresivo de la médula ósea (Alter, 2003);

no obstante, debido a la alta heterogeneidad característica de la enfermedad, no todos estos síntomas se presentan en todos los pacientes. Entre las diferentes manifestaciones clínicas y complicaciones pueden aparecer (Shimamura & Alter, 2010):

- Fallo progresivo de la médula ósea. Anemia aplásica (pancitopenia).
- Susceptibilidad a la aparición de leucemia, cáncer o tumores sólidos.
- Retraso mental, complicaciones en el sistema nervioso central.
- Microcefalia.
- Pérdida de capacidad auditiva.
- Malformaciones esqueléticas.
- Retraso del crecimiento/desarrollo, baja estatura.
- Pigmentación anormal: hipopigmentación, hiperpigmentación o manchas café con leche.
- Fallo renal.
- Hipogonadismo.
- Complicaciones cardiopulmonares.
- Complicaciones gastrointestinales.

Si nos centramos en las complicaciones neoplásicas, cerca del 10% de los pacientes con FA acaban desarrollando leucemias, principalmente leucemia mieloide aguda, y el 7% desarrolla síndrome mielodisplásico (Alter, 2003). Alrededor del 5% de los pacientes desarrolla algún tipo de tumor sólido y cerca de un 3% tumores en hígado. El porcentaje de fallecidos con algún tipo de complicación neoplásica

se sitúa muy por encima de los pacientes en los que no aparecen estas complicaciones (tabla 5).

Características	Todos los casos	Leucemia	Síndrome mielodisplásico	Tumor sólido	Tumor en hígado
Nº de casos	1301	116	89	68	37
Porcentaje del total	100	8.9%	6.8%	5.3%	2.8%
Hombre/mujer	711:578	70:46	46:42	23:45 (23:34) ^a	22:15
Ratio	1.2	1.5	1.1	0.5 (0.59) ^a	1.5
Edad de diagnóstico					
Media	8.3	10.1	11.3	12.7	9.2
Mediana	7	8.6	9.3	9	7
Rango de edad	0-48	0.13-28	0.2-43	0.44	3-48
Edad de aparición de la complicación					
Media	-	14.5	15.7	22.6	15.7
Mediana	-	14	14	25.5	13
Rango de edad	-	0.13-29	1.8-43	0.2-45	6-48
Nº de fallecidos	488	84	44	41	30
Porcentaje de fallecidos	38%	72%	51%	61%	81%
Edad de supervivencia media estimada	20	16	21	31	16

Tabla 5: Casos clínicos de anemia de Fanconi, características y complicaciones. Datos recopilados desde 1927 hasta 2001. a. Excluyendo cáncer de tipo ginecológico. Adaptado de Alter, 2003.

El fallo progresivo de la médula ósea se suele presentar con anemia aplásica en la primera década de vida, pudiendo presentarse pancitopenia incluso en el momento del nacimiento (Shimamura & Alter, 2010). La anemia aplásica es uno de los síntomas más comunes entre los pacientes FA, de hecho se estima que la incidencia de fallo en la médula ósea es del 90% para pacientes de FA de entre 40 y 50 años.

Los pacientes de FA presentan un riesgo relativo de desarrollar AML unas 500 veces superior al de individuos sanos (Alter, 2003). Asimismo, se ha demostrado que el riesgo de desarrollar tumores sólidos se encuentra aumentado particularmente en cuello, piel, cabeza, esófago y zonas ginecológicas (tabla 6) (Kutler *et al.* 2004, Rosenberg *et al.* 2003,

Rosenberg *et al.* 2008, Alter *et al.* 2010, Tamary *et al.* 2010).

Cáncer	Edad en población (a)	Edad en FA	Rango de edad en FA	Nº en FA
Cualquiera	68	16	0.1-48	211
Cualquier tumor sólido	-	26	0.2-45	68
Leucemia (AML)	68	14	0.1-29	116
Hígado	68	13	6-48	37
Cuello y cabeza	64	28	13-41	26
Cuello y cabeza (b)	-	21	11-33	12
Esófago	68	27	20-36	9
Vulva/Ano	72	27	20-37	10
Cervix	47	25	22-32	3
Cerebro	56	3	0.5-11	6
Mama	63	37	26-45	4

Tabla 6: Complicaciones neoplásicas en anemia de Fanconi. Datos recopilados desde 1927 hasta 2001. a. Edad media de aparición en la población general. b. Tras trasplante de médula ósea (*Bone Marrow transplant*, BMT). Adaptado de Alter, 2003.

1.3.3.- Diagnóstico y tratamiento.

Dada su elevada heterogeneidad y las diferencias en la edad de aparición de algunos síntomas, en muchas ocasiones el diagnóstico de la enfermedad resulta complicado; en algunos casos, incluso, se produce tras el reconocimiento de alguna complicación maligna, como ocurre con algunos pacientes que no presentan fallo en la médula ósea, pero que acaban desarrollando tumores sólidos (revisado en Alter, 2003). Destaca la correlación entre la aparición tardía de estos tumores con la ausencia de otros síntomas hematológicos (tabla 6).

A nivel de pruebas diagnósticas de la enfermedad, el test clásico se basa en la detección de alteraciones cromosómicas en las células de los

pacientes, tras someterlas a cultivo con agentes intercalantes del DNA tales como la mitomicina C (MMC) o el diepoxibutano (DEB) (Auerbach, 1993). Tras el tratamiento con agentes intercalantes del DNA, la aberración cromosómica más característica de la anemia de Fanconi es la aparición de cromosomas radiales (Figura 16); no obstante, recientemente se ha descrito que es más informativo medir el porcentaje de aberraciones por célula que el número de aberraciones cromosómicas por célula, ya que existen otros síndromes, como el síndrome de Roberts o el de rotura de Nijmegen, que también presentan sensibilidad a agentes intercalantes de DNA (van der Lelij *et al.*, 2010).

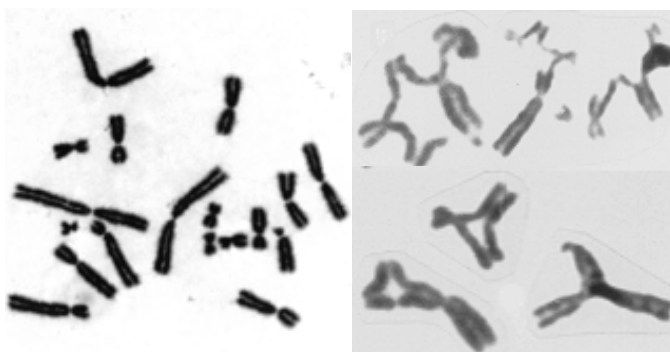


Figura 16. Cromosomas radiales: estas alteraciones cromosómicas (derecha) aparecen cuando se expone las células de pacientes afectados a dosis bajas de agentes intercalantes de DNA, mientras que en células no afectadas los cromosomas conservan su morfología normal (izquierda). Se utilizan como criterio diagnóstico para detectar la enfermedad, suponiendo el rasgo más característico de los cariotipos de las células de Anemia de Fanconi.

Una vez se confirma la sensibilidad al DEB o la MMC, el diagnóstico pasa por identificar, mediante complementación con retrovirus (Chandra *et al.*, 2005), el gen mutado en el paciente, tras lo cual se puede pasar a analizar la existencia de variantes de la secuencia del gen FANC o posibles duplicaciones o deleciones de la región mediante secuenciación (Ameziane *et al.*, 2009).

En cuanto al tratamiento, actualmente no existe ninguna terapia curativa viable para todos los pacientes; sin embargo, lo que se suele hacer es realizar un trasplante de médula ósea (BMT) de donantes con un HLA idéntico (normalmente de un familiar), mediante trasplante de células progenitoras hematopoyéticas (HSCT). Estos trasplantes pueden curar las complicaciones como la anemia o el síndrome mielodisplásico; no obstante, es preferible realizar el trasplante antes de la aparición de dichas complicaciones (MacMillan & Wagner, 2010). Este tratamiento es curativo en cuanto a los síntomas hematológicos, en los pacientes en los que no se produce rechazo, pero no previene ni cura otras complicaciones que pueden aparecer, ya que es frecuente encontrar pacientes que acaban desarrollando tumores sólidos tras someterse a un trasplante de médula ósea (revisado en Alter, 2003; Rosemberg *et al.*, 2005).

Clásicamente, para el tratamiento se utilizan andrógenos para incrementar el recuento de células sanguíneas aunque, recientemente, se ha descrito que su uso puede acarrear efectos adversos, destacando la aparición de daños en el hígado de diversa índole, pudiendo incluso desarrollarse tumores hepáticos (Shimamura & Alter, 2010).

Dada la existencia de mosaicismo celular en muchos pacientes, y aunque infrecuente, existe la posibilidad de que se produzca una reversión genotípica, que desemboca en una corrección fenotípica, es decir, la enfermedad puede remitir cuando se corrige el defecto causante en una célula progenitora, que con el tiempo y debido a que posee ventajas en cuanto a viabilidad y proliferación, puede acabar dominando la población hematopoyética de la médula ósea, dando como resultado la reversión del mosaicismo (Waisfisz *et al.*, 1999); si bien, esto no implica necesariamente la desaparición de la inestabilidad genómica en otras células del paciente, o la corrección del riesgo de desarrollar tumores.

La enorme complejidad que presenta el tratamiento de esta enfermedad, se pone de manifiesto a la hora de seleccionar un método adecuado para tratar los tumores o las leucemias que pueden desarrollar estos pacientes. Esto se debe a que la mayoría de los tratamientos antitumorales están basados en producir daños en el DNA, para interferir con la replicación y así controlar la proliferación desmesurada de las células cancerígenas. En el caso de la anemia de Fanconi, las células “sanas” presentan dificultades para la proliferación (incluso sin tratamiento alguno), y una extrema sensibilidad a los agentes intercalantes del DNA, tales como la Mitomicina C o el Cisplatino. Todos estos motivos complican enormemente la elección de un tratamiento adecuado ante el desarrollo de tumores, especialmente teniendo en cuenta que la mayoría de los agentes quimioterápicos que se utilizan en la clínica son concretamente agentes intercalantes del DNA.

1.3.4.- Etiología y Bases Moleculares: ruta de la Anemia de Fanconi.

La aparición de las diferentes manifestaciones clínicas de la enfermedad tiene su base genética en una coordinación incorrecta de la reparación y/o la replicación en las células afectadas. La descoordinación de las diferentes rutas implicadas se produce como consecuencia de mutaciones en la ruta de la Anemia de Fanconi (*FA pathway* o **FA/BRCA pathway**), que se sitúa en una rama central como reguladora de la reparación y/o replicación de regiones afectadas por diferentes tipos de lesiones. A grandes rasgos, el papel fundamental que juega la ruta FA es principalmente la coordinación de la recombinación homóloga (o **HR** de *Homologous Recombination*) y la síntesis translesión (o **TLS** de *TransLesion synthesis*), pero tales vías guardan una estrecha relación entre sí, así como con otras rutas de reparación, como son la vía de escisión de nucleótidos (o **NER** por *Nucleotide Excision Repair*), la unión de DNA de cadena sencilla (o **SSA** por *Single Strand Annealing*) y la vía de unión de extremos no homólogos (o **NHEJ** por *Non-Homologous End Joining*) (Revisado en Moldovan & D'andrea, 2009).

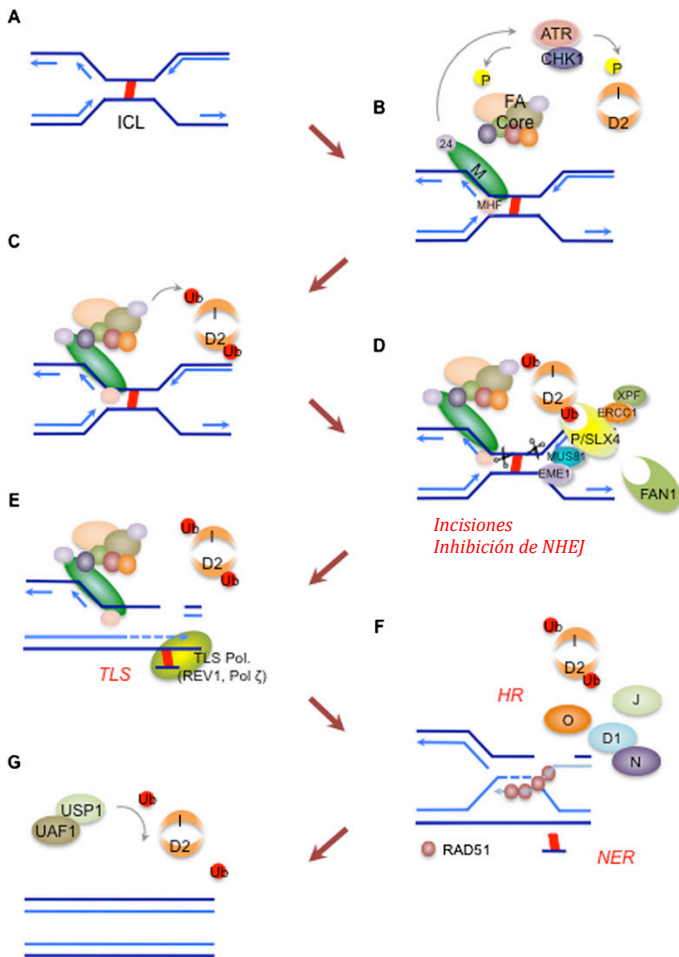


Figura 17. Ruta de reparación de Anemia de Fanconi (FA/BRCA pathway). La ruta coordina los diferentes procesos durante la reparación de los ICLs. A, atasco de las horquillas de replicación; B, reclutamiento del complejo FANCD1/FANCD2 (FA Core); C, ubiquitinación de FANCD2 y FANCD1; D, incisiones a ambos lados de la ICL; E, síntesis translesión; F, recombinación homóloga y escisión de nucleótidos dañados; G, liberación de FANCD2 y FANCD1.

A pesar de que aún hay muchas preguntas que responder acerca del funcionamiento y mecanismos implicados en esta ruta; lo que sí se sabe a ciencia cierta, es que esta vía es imprescindible para reparar eficientemente un tipo concreto de daño que se produce en el DNA, los cruzamientos intercatenarios (*Inter Strand Crosslinks* o ICLs), que aparecen cuando una de las cadenas (o hebras) de DNA se une de manera covalente a la otra (revisado en Kim & D'Andrea, 2012). Clásicamente, éste es el papel fundamental de la ruta de la anemia de Fanconi, debido a la enorme cantidad de procesos que se deben suceder para que la célula repare correctamente estas lesiones; sin embargo, cada día aparecen nuevas evidencias que apuntan a que la ruta juega un papel mucho más amplio en la replicación y reparación del DNA, siendo fundamental para dar una respuesta adecuada a la aparición de diversas fuentes de daño, tanto internos como externos, además de los ICLs (figura 17) (Sharma *et al.*, 2011; Fu *et al.*, 2013; revisado en Kee & D'andrea, 2012; Kottemann & Smogorzewska, 2013).

No obstante, es de gran importancia el procesamiento correcto y reparación de las ICLs, ya que este tipo de lesiones presentan una elevada toxicidad para la célula, debido a que bloquean las maquinarias de replicación y transcripción, pudiendo llevar a la muerte de la célula (revisado en Deans & West, 2011). Asimismo, la reparación incorrecta de este tipo de daño parece ser causa directa de inestabilidad genómica, como demuestra la aparición de los cromosomas radiales y otras aberraciones cromosómicas tras someter las células a dosis bajas de agentes intercalantes de DNA (revisado en Deans

& West, 2011). Otra prueba de este fenómeno es que paradójicamente, tanto la aparición de aberraciones como la inestabilidad genómica, pueden ser corregidas, en mayor o menor medida, suprimiendo la vía de reparación de unión de extremos no homólogos (NHEJ), que repararía de manera menos fidedigna las estructuras intermediarias que se producen durante la reparación del daño causado por los agentes intercalantes (Pace *et al.*, 2010; Adamo *et al.*, 2010).

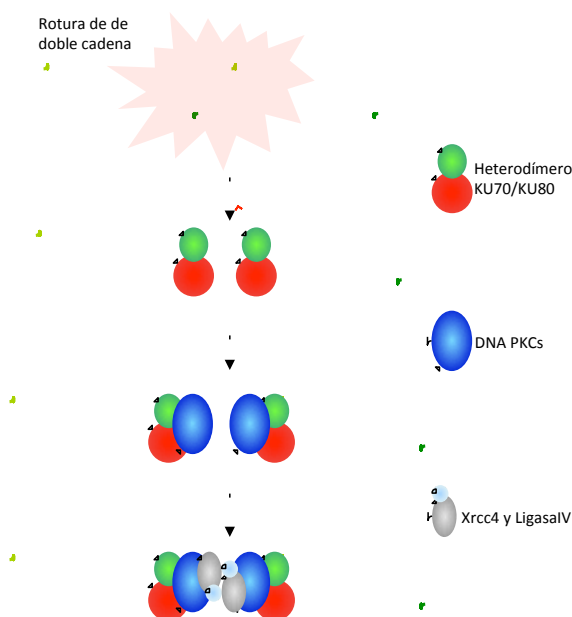


Figura 17. Reparación de roturas de doble cadena mediante la unión de extremos no homólogos (NHEJ). La maquinaria de reparación recluta al heterodímero Ku70/Ku80, que es reconocido por las proteín-kinasas C (PKCs), y finalmente por la proteína XRCC4 y la ligasa IV, que unen la doble cadena.

1.3.5.- Miembros de la ruta de Anemia de Fanconi.

Los miembros de la ruta de FA también se conocen como grupos de complementación. La mayoría de las proteínas de la ruta FA (tabla 7) no contiene ningún dominio enzimático reconocible, lo que sugiere que actúan a modo de soporte o andamio. El complejo se presenta en la figura 18.

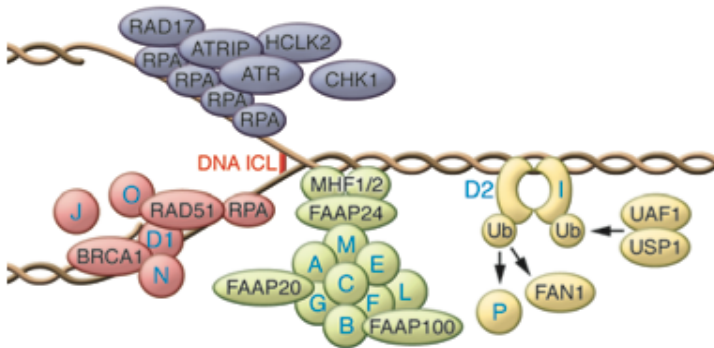


Figura 18. Miembros de la ruta de anemia de Fanconi (FA/BRCA pathway). La ruta FA es la única que puede reparar de forma eficiente los cruzamientos entre las dos cadenas de DNA (Interstrand Cross Links, DNA ICLs). Con rotulos en negro se indican proteínas de otras rutas que no forman parte de la vía pero que participan con la ruta FA en el proceso de reparación de ICLs. Adaptado de Kee & D'andrea, 2012.

Si los agrupamos en cuanto a su función distinguimos principalmente 4 grupos (revisado en Kee & D'andrea, 2012; Kottemann & Smogorzewska, 2013):

- Los que forman el **complejo principal FA**: FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL y FANCM, así como las proteínas asociadas FAAP20, FAAP24 Y FAAP100. Forman un complejo nuclear, denominado complejo principal FA, que posee una actividad E3 ubiquitin-ligasa, la función de este complejo es la ubiquitinación del complejo efector D2/I.

- Los que participan en la coordinación de la **recombinación homóloga**: BRCA2 (FANCD1), PALB2 (FANCN), RAD51C (FANCO). Los miembros de este grupo parecen no afectar a la ubiquitinación de FANCD2.

- Complejo principal **ID**: compuesto por FANCD2 y FANCI, es considerado el efector principal de la ruta. FANCD2 y FANCI son expresadas de manera constitutiva, se encuentran unidas al DNA como un heterodímero, su ubiquitinación actúa como señal para reclutar otras proteínas de reparación.

- Factores no necesarios para la ubiquitinación de D2: BRIP1 (FANCF), SLX4 (FANCP), FAN1 (*Fanconi Associated Nuclease 1*). A pesar de no formar parte del complejo principal ni participar directamente en la recombinación homóloga, la actividad de estos factores es necesaria para el correcto funcionamiento del resto de los componentes de la vía.

- Existe también un complejo accesorio, formado por BRCA2, FANCD2, FANCG y XRCC3, llamado **D1-D2-G-X3**, de función desconocida pero que parece vincular

la ruta FA con la recombinación homóloga y es independiente de la ubiquitinación de FANCD2, se requiere la participación de ATR para su formación (Wilson *et al.*, 2010).

A modo de resumen en la tabla 7 se muestra la prevalencia de cada uno de los componentes y su localización cromosómica.

GEN	PREVALENCIA APROXIMADA	LOCALIZACIÓN	TAMAÑO DE LA PROTEINA (kDa)
FANCA	66-70%	16q24.3	163
FANCB	2%	Xp22.31	95
FANCC	10-14%	9q22.3	63
FANCD1 (BRCA2)	2-3%	13q12-13	380
FANCD2	2-3%	3q25.3	155, 162
FANCE	2-3%	6p21-22	60
FANCF	2%	11p15	42
FANCG	9-10%	9p13	68
FANCI	1-2%	15q25-26	140, 147
FANCI (BRIP1)	2%	17q22-24	140
FACNL	0,2%	2p16.1	43
FANCM	0,2%	14q21.3	250
FANCN (PALB2)	0,7-2%	16p12.1	140
FANCO (RAD51C)	0,2%	17q22	42
FANCP (SLX4)	0,2%	16p13.3	200

Tabla 7: Genes de la ruta de Anemia de Fanconi y su localización cromosómica. Se indica la prevalencia de cada grupo de complementación y el tamaño de la proteína. Adaptado de Moldovan & D'andrea, 2009; Shimamura & Alter, 2010.

1.3.5.1.- FANCA.

Es, con diferencia, el miembro de la vía cuyas mutaciones aparecen con mayor frecuencia en los pacientes, constituyendo aproximadamente un 66% de los casos diagnosticados de FA, y en la población española la proporción alcanza el 80% (Castella *et al.*, 2011; revisado en Moldovan & D'andrea, 2009).

Se localiza en el cromosoma 16, concretamente en la citobanda 16q24.3. Esta proteína presenta principalmente dos isoformas, una más larga, denominada isoforma a, con 43 exones, codifica una proteína de 1455 aminoácidos, que forma parte del complejo principal FA (Apostolou *et al.*, 1996). FANCA es una fosfoproteína que es fosforilada en su serina 1449 por la acción de ATR (Collins *et al.* 2009). Esta proteína contiene dos señales de localización nuclear (NLS), cinco señales de exportación nuclear (NES) ricas en leucina, y una secuencia parcial de cremallera de leucina (Ferrer *et al.*, 2005).

1.3.5.2.- FANCB.

También llamada FAAP95, las mutaciones en este gen suponen un 2% de los casos diagnosticados y su localización genómica es en el cromosoma X, Xp22.2. Posee 10 exones, aunque su traducción empieza en el exon 3, codificando una proteína de 853 aminoácidos. Forma parte del complejo principal FA y contiene una supuesta señal NLS. FANCB es el único grupo de complementación cuya herencia está ligada al cromosoma X (Meetei *et al.*, 2004).

1.3.5.3.- FANCC.

Representa el segundo grupo de complementación más frecuente en pacientes de FA, con un 10-14% de los casos diagnosticados (Moldovan & D'andrea, 2009; Shimamura & Alter, 2010).

Se localiza en el cromosoma 9, la citobanda 9q22.32, y presenta 15 exones que codifican una proteína de 558 aminoácidos. Forma parte del complejo principal FA, pero su localización parece ser tanto nuclear como citoplasmática (Yamashita *et al.* 1994). Se ha sugerido que FANCC puede tener otras funciones fuera de su participación en la vía clásica de FA (Fagerlie *et al.* 2004).

1.3.5.4.- FANCD1/BRCA2.

Aproximadamente un 3% de los casos de FA se deben a mutaciones en el gen BRCA2 (de *BReast CAncer type 2 susceptibility protein*), siendo el gen de la ruta que parece presentar mayor proporción de letalidad embrionaria. Las mutaciones bialélicas en BRCA2 son escasas (Magdalena *et al.*, 2005; Castella *et al.*, 2011; Shimamura & Alter, 2010), pero cuando se producen destaca la aparición de leucemias agudas y tumores sólidos a muy temprana edad, siendo el riesgo relativo de padecer algún tipo de malignización del 97% a la edad de 6 años (Alter *et al.*, 2007). No obstante, a pesar de que en la anemia de Fanconi las mutaciones suelen ser bialélicas y por lo tanto la enfermedad recesiva, se ha descrito que, los individuos heterocigotos para una mutación en BRCA2, no presentan la sintomatología de la anemia

de Fanconi y son individuos sanos, pero presentan un riesgo elevado de desarrollar cáncer de mama (Howlett *et al.*, 2002).

Su localización cromosómica es 13q13.1, está compuesto por 27 exones que codifican una proteína de 3418 aminoácidos; BRCA2 no forma parte del complejo principal FA y no es necesario para la ubiquitinación de FANCD2.

La función principal de BRCA2 es regular la reparación de roturas de doble cadena DNA por recombinación homóloga mediante la recombinasa RAD51 (Davies *et al.*, 2001); no obstante, BRCA2 presenta otras funciones conocidas, como son la estabilización de las horquillas de replicación y la regulación de la citocinesis (Daniels *et al.*, 2004).

BRCA2 interactúa con BRCA1 en la recombinación homóloga y dicha interacción parece estar mediada por PALB2 (FANCN) (Xia *et al.*, 2007), mostrándose una vez más la estrecha relación y la comunicación cruzada entre las vías FA y de recombinación homóloga. El papel de BRCA1 parece ser promover la carga y asentamiento de RAD51 sobre el ssDNA, BRCA1 tiene efectos sobre la cromatina, y además llega a las roturas en el DNA independientemente de la actuación de H2AX (revisado en Nagaraju & Scully, 2007).

Tanto BRCA1 como BRCA2 son indispensables para el correcto desarrollo embrionario, como se demostró en embriones de ratón, donde la ausencia de uno de los dos provoca letalidad embrionaria en un período muy corto de tiempo, en el que además el embrión se ve sometido a un alto estrés, revelado por una elevada expresión de p53 y p21; no obstante, parecen tener efectos epistáticos, ya que la ausencia

de ambos provoca los mismos efectos en el mismo tiempo de desarrollo embrionario, aunque esta letalidad puede ser ligeramente pospuesta si además se elimina p53 (revisado en Nagaraju & Scully, 2007).

BRCA2 interacciona también con FANCD2 y dicha interacción está mediada por FANCG y XRCC3 y no requiere la ubiquitinación de FANCD2 (Wilson *et al.*, 2008, 2010). Asimismo se ha demostrado la colocalización de FANCD2 ubiquitinado y BRCA2 en cromatina, formando un complejo con FANCE (Wang *et al.*, 2007).

BRCA2 es un gen supresor de tumores que parece jugar un papel dual en la respuesta al daño en el DNA: por un lado, se encarga de regular la recombinación homóloga, mientras que por otro interacciona de manera directa con FANCD2; esta interacción entre ambas proteínas parece ser clave para la coordinación de los eventos que se deben suceder en el DNA para la reparación correcta del daño detectado, quizá actuando a modo de interruptor entre la síntesis translesión y la recombinación homóloga.

1.3.5.5.- FANCD2.

Las mutaciones bialélicas en FANCD2 son escasas, representando un 3% de los casos (Shimamura & Alter, 2010). Existen dos isoformas, la isoforma *b* codifica una secuencia más corta, que contiene 43 exones y parece no ser funcional, mientras que la isoforma *a* está compuesta por 44 exones y una secuencia en 3' UTR alternativa, que da lugar a una proteína más corta y con un dominio C-

terminal distinto, que es la forma funcional de la proteína (Montes de Oca *et al.*, 2005).

Su localización genómica es en la citobanda 3q25.3. La proteína funcional está compuesta por 1451 aminoácidos, tiene un peso molecular de 162kDa y posee cierta similitud con FANCI. Ambas proteínas forman el complejo ID, que se considera el efector principal de la ruta FA. FANCD2 es ubiquitinado en su lisina 561 mediante la participación del complejo principal (García-Higuera *et al.*, 2001), concretamente por FANCL, de una forma dependiente de UBE2T y FANCI (Smogorzewska *et al.*, 2007). La forma monoubiquitinada de FANCD2 forma foci nucleares que colocalizan con otras proteínas como FANCI, BRCA1, BRCA2 o RAD51.

FANCD2 es fosforilado por ATM (por *Ataxia Telangiectasia Mutated*) (Taniguchi *et al.*, 2002), y posiblemente por ATR (por *Ataxia Telangiectasia and Rad3-related protein*) (Andreassen *et al.*, 2004, Pichierri & Rosselli 2004), cuando se produce daño en el DNA. Se ha demostrado que para unirse a BRCA2, FANCD2 tiene que ser fosforilada en su serina 331 por CHK1 (Zhi *et al.*, 2009), y otros autores demuestran que dicha interacción requiere además la fosforilación de la serina 7 de FANCG (por ATR o CHK1), mediaría la interacción directa de BRCA2 y FANCD2, junto con XRCC3 y el propio FANCG (Wilson *et al.*, 2008). Además se ha demostrado que FANCD2 es capaz de unirse a la proteína conocida como PCNA (*Proliferative Cell Nuclear Antigen*) (Howlett *et al.*, 2009). Se ha sugerido que esta interacción es necesaria para reclutar la polimerasa adecuada para realizar la síntesis translesión, en función del tipo de daño reconocido; de hecho, se ha demostrado que

una de estas polimerasas, la polimerasa “*eta*”, interacciona antes con FANCD2 que con PCNA, en células expuestas a radiación ultravioleta, haciéndolo a través de los mismos dominios que median la interacción de esta polimerasa con PCNA (Fu *et al.*, 2013).

Otra de las interacciones claves para el desarrollo correcto de la vía es la interacción que se produce entre FANCD2 y FAN1, una nucleasa asociada a la ruta de Fanconi (Huang & D’Andrea, 2010); asimismo, se ha demostrado que FANCD2 participa en el procesamiento de los extremos de las roturas que se producen durante la reparación de los cruzamientos intercatenarios (ICLs) del DNA (Knipscheer *et al.*, 2009).

Recientemente se ha demostrado además la participación de FANCD2 ubiquitinado, de manera conjunta con FANCI, en la reestructuración de la cromatina mediante su interacción con histonas, facilitando el correcto ensamblaje de los nucleosomas ante daños en el DNA. Se ha sugerido que esta función está relacionada con el reclutamiento adecuado de nucleasas para reparar la lesión, previniendo cortes sobre el DNA de cadena sencilla (Sato *et al.*, 2012).

1.3.5.6.- FANCE.

Aproximadamente un 2% de los casos de FA se deben a mutaciones en este gen, localizado en la posición cromosómica 6p21-22, y que posee 14 exones que codifican una proteína de 536 aminoácidos y 60kDa. También forma parte del complejo principal FA, pero parece unirse directamente a FANCD2. FANCE contiene dos señales

NLS y cinco repeticiones en tándem de un dominio hélice corto (Nookala *et al.*, 2007). Parece actuar a modo de proteína lanzadera entre el complejo principal FA y FANCD2, de una manera dependiente de su fosforilación; FANCE también se une a FANCC (Wang *et al.*, 2007). Asimismo, se ha sugerido que puede participar en la localización en la cromatina de BRCA2, miembro con el que interactúa de manera independiente de FANCD2 y del complejo D1-D2-G-X3 (Wang *et al.*, 2007).

1.3.5.7.- FANCF.

Representa un 2% de los casos de FA; con tan sólo un exón, codifica una proteína de 374 aminoácidos y 42kDa que forma parte del complejo principal FA. FANCF parece actuar a modo de adaptador para el correcto ensamblaje del complejo FA (Léveillé *et al.*, 2004).

1.3.5.8.- FANCG.

Las mutaciones bialélicas en este gen suponen un 10-14% de todos los casos de FA. El gen contiene 14 exones que codifican una proteína de 622 aminoácidos y 68kDa. FANCG forma parte del complejo principal FA, pero parece funcionar también en la recombinación homóloga, promoviendo la formación del complejo D1-D2-G-X3, constituyendo por tanto una vía de comunicación de la ruta FA con la recombinación homóloga. FANCG está compuesta por siete motivos TPR (*Tetratricopeptide Repeat*) (Blom *et al.*, 2004), por los que parece interactuar con XRCC3, BRCA2 (Hussain *et al.*, 2006) y FANCD2

(Wilson *et al.*, 2010). Estos motivos TPR son también responsables de su interacción con ERCC1-XPF, una endonucleasa fundamental para la reparación de los daños causados por las ICLs (Wang & Lambert, 2010).

FANCG es una fosfoproteína que se fosforila en sus serinas 383 y 387, presumiblemente por *cdc2* durante la fase M del ciclo celular (Mi *et al.*, 2004). Estos dos sitios de fosforilación parecen importantes para separar a FANCG de la cromatina durante la mitosis; por otra parte, se ha descrito también que la fosforilación de esta proteína en su serina 7 aumenta tras exposición a MMC (Qiao *et al.*, 2004), y que dicha fosforilación es fundamental para la interacción entre BRCA2 y FANCD2, así como para la interacción de FANCG con XRCC3, BRCA2 y FANCD2 y por tanto para la formación del complejo D1-D2-G-X3; sin embargo, la fosforilación en la serina 7 no es necesaria para la ubiquitinación de FANCD2 (Wilson *et al.*, 2008).

1.3.5.9 FANCI.

Las mutaciones bialélicas en FANCI causantes de la enfermedad son muy escasas, constituyendo tan sólo un 1-2% de los casos de FA. Este gen posee 37 exones que codifican una proteína de 1328 aminoácidos y 140kDa, y que tiene similitud con FANCD2 (Smogorzewska *et al.*, 2007); FANCD2 y FANCI ubiquitinados, de manera interdependiente, componen el complejo ID. FANCI es monoubiquitinado en su lisina 523, para ello se requiere la acción del complejo principal FA, UBE2T y FANCD2. FANCI es una fosfoproteína que se fosforila en la serina730, la treonina952 y la serina1121,

presumiblemente mediante la acción de ATM y/o ATR; dicha fosforilación se produce tras daño en el DNA (Smogorzewska *et al.*, 2007). Su función parece ser análoga a la de FANCD2.

1.3.5.10 FANCJ.

También conocido como Brip1 (*BRCA1 interacting protein C-terminal helicase 1*), constituye el 2% de los casos de FA y está compuesto por 20 exones que codifican una proteína de 1249 aminoácidos y 140kDa.

FANCJ es una helicasa 5'-3' con actividad ATPasa, que se une directamente al dominio BRCT de BRCA1 (Cantor *et al.*, 2001), y que no es necesaria para la ubiquitinación del complejo ID. Curiosamente, a pesar de su interacción con BRCA1, la ausencia de FANCJ provoca la sintomatología propia de la FA, demostrando así su implicación en la ruta de reparación por recombinación homóloga, y al igual que ocurre con BRCA2, pone de manifiesto de nuevo la importancia de que exista una perfecta coordinación de ambas rutas. Se ha demostrado que los niveles de expresión de FANCJ controlan los niveles de expresión de BLM, otra helicasa que participa en la reparación del DNA y cuya alteración es responsable de la aparición del síndrome de Bloom; de esta manera, los niveles de expresión de BLM caen drásticamente en células deficientes en FANCJ (Suhasini *et al.*, 2011). Ambas helicasas, que no colocalizan en condiciones normales, lo hacen en situaciones de estrés replicativo inducido por tratamiento con hidroximetilglutimidilglutaminato; sin embargo, hay que

destacar que los niveles de expresión de BLM no afectan a FANCL (Suhasini *et al.*, 2011).

1.3.5.11.- FANCL.

Las mutaciones en este gen parecen ser extremadamente escasas en la clínica, representando tan sólo un 0,2% de los casos diagnosticados de FA. FANCL está compuesto por 14 exones, codifica una proteína de 375 aminoácidos y 43 kDa, con función E3 ubiquitina-ligasa. FANCL forma parte del complejo principal FA, considerándose como la subunidad catalítica de éste, ya que es la encargada de la ubiquitinación de las proteínas del complejo ID.

FANCL posee tres dominios WD40 y un motivo de dedo PHD (una variedad de los dedos RING) (Meetei *et al.*, 2003). Esta proteína interacciona directamente con UBE2T (enzima conjugadora de ubiquitina E2) y ubiquitina a FANCD2 y FANCI (Machida *et al.*, 2006, Smogorzewska *et al.*, 2007).

1.3.5.12.- FANCM.

Las mutaciones en este gen son también muy escasas en la clínica, representando tan sólo un 0,2% de los casos. FANCM está formado por 23 exones que codifican una proteína de 2048 aminoácidos y 250 kDa. FANCM actúa como translocasa formando parte del complejo principal FA, contiene siete motivos específicos de helicasa, un dominio endonucleasa degenerado, y posee actividad ATPasa estimulada por DNA de cadena doble o sencilla (Meetei *et al.*, 2005). FANCM es fosforilada cuando se produce daño en el DNA y, junto con la proteína FAAP24, participa en la

reacción de *checkpoint* de respuesta a daño en el DNA (Huang *et al.*, 2010).

1.3.5.13.- FANCN.

Las mutaciones en este gen, también conocido como Palb2 (*Partner And Localizer of BRCA2*), se encuentran en un 0,7-2% de los pacientes de FA. Contiene 13 exones que codifican una proteína de 1186 aminoácidos y 140 kDa, cuya función principal parece ser regular la localización y estabilidad de la proteína BRCA2. FANCN no es necesaria para la ubiquitinación de FANCD2 y, junto con FANCI y BRCA2 parece jugar un papel muy importante en la recombinación homóloga (Xia *et al.*, 2007), concretamente favoreciendo la interacción entre BRCA2 y BRCA1; de hecho, es uno de los genes de la ruta de anemia de Fanconi que es también marcador de susceptibilidad a cáncer de mama.

1.3.5.14.- FANCO.

Más conocido como RAD51C, presenta una prevalencia aproximada de un 0,2% en pacientes de FA. Este gen contiene tan sólo 9 exones que codifican una proteína de 376 aminoácidos y 42kDa, que participa en complejos implicados en la recombinación homóloga. FANCO es uno de los parálogos de RAD51, que poseen la capacidad de unirse a extremos de DNA de cadena sencilla, que suelen producirse producen tras el procesamiento de lesiones en el DNA (Somyajit *et al.*, 2010). RAD51C interacciona con XRCC3 y promueve su fosforilación.

Es también uno de los genes que determinan susceptibilidad a cáncer de mama.

1.3.5.15.- FANCP.

También conocida como SLX4, supone un 0,2% de los casos de FA y tiene 15 exones que codifican una proteína de 1834 aminoácidos y 200kDa, posiblemente implicada en la resolución de estructuras intermediarias de la recombinación homóloga, como las uniones de Holliday (*Holliday junctions*). Se ha propuesto que esta endonucleasa interacciona y coordina la actuación de otros complejos de endonucleasa, concretamente SLX1, MUS81-EME1 y XPF-ERCC1 (Kim *et al.*, 2001, Stoepker *et al.*, 2011). Asimismo, se ha visto que interacciona con la helicasa BLM.

1.3.5.16.- FAN1.

La nucleasa asociada a fanconi FAN1 (*Fanconi-Associated Nuclease 1*) no es propiamente un gen FA, ya que no se han descrito pacientes con mutaciones en FAN1 que presenten la sintomatología típica de FA; sin embargo, participa en la reparación de las ICLs reclutado por FANCD2 monoubiquitinado, mediante sus dominios UBZ de unión a ubiquitina, y FANCD2 coordina su actividad nucleasa para efectuar incisiones en el DNA próximas a las ICLs (Huang *et al.*, 2010).

1.3.6.- Reparación de cruzamientos intercatenarios (ICLs)

La amenaza que presentan los cruzamientos intercatenarios o ICLs (por *Interstrand CrossLinks*) durante la interfase es menor que durante la replicación, ya que el bloqueo de las RNA polimerasas generalmente no conduce a la muerte de la célula; sin embargo, el atasco de las polimerasas durante la replicación provoca paradas en las horquillas de replicación, lo que puede llevar al colapso de éstas, al arresto del ciclo celular e incluso a la muerte programada de la célula.

Así, uno de los puntos principales a tener en cuenta, desde el punto de vista celular, es la fase en la que aparecen las ICLs, ya que durante su reparación es necesaria la generación de roturas de doble cadena (DSBs) a modo de lesiones intermedias; ahora bien, en condiciones normales, la ruta que repara las DSBs está mayormente controlada por el momento del ciclo en el que se encuentra la célula, y en la interfase (fase G1), predomina la vía de unión de extremos homólogos (*Non-Homologous End Joining* o NHEJ) sobre la HR, mientras que durante la replicación (fase S), predomina la HR sobre la NHEJ (revisado en Branzei & Foiani, 2010).

Si pensamos en la interfase, la reparación de las DSBs mediante la NHEJ (figuras 17 y 19) se presenta como la vía más adecuada para reparar eficientemente las DSBs, por ser una manera más rápida de reparar el daño que la HR, y seguramente menos propensa a cometer errores durante la interfase que durante la replicación, ya que esta vía

una o dos roturas que se encuentran próximas entre sí. Así, para que la NHEJ introduzca un error durante la reparación en la interfase celular, es necesario que se generen dos roturas que generen DSBs en dos sitios distintos a la vez (un total de 4 DSBs) y que estas roturas se encuentren próximas entre sí.

En cambio, durante la replicación, la generación de DSBs es potencialmente problemática, ya que, junto a una cadena doble de DNA se encuentra otra cadena doble de DNA, en paralelo, y una tercera doble cadena, aguas abajo de las dos dobles cadenas nuevas (si tenemos en cuenta la cadena original aún no separada). Las dos cadenas que se van generando están separadas únicamente por la acción de helicasas, y aquí la posibilidad de que se produzcan dos roturas en dos sitios cercanos a la vez es mayor. A esto hay que añadir la presencia constante de fragmentos de DNA de cadena sencilla producidos por el proceso de replicación discontinuo de la cadena retardada, debido a su vez al mecanismo de replicación continuo de la cadena líder (figura 19).

Dado que la NHEJ es un mecanismo menos específico, y por lo tanto más propenso a la aparición de errores, constituye la causa principal de las aberraciones cromosómicas detectadas en las células FA. Esta explicación está respaldada por el descubrimiento, en varias líneas celulares, de que al interferir de alguna manera con la vía NHEJ en las células FA, se consigue rescatar parcial, e incluso totalmente, la sensibilidad a los agentes intercalantes de DNA (revisado en Kottemann y Smogorzewska, 2013).

Por tanto, durante la replicación del DNA es más seguro para la célula que predomine la HR, para reparar eficientemente las DSBs generadas a partir de las ICLs; sin embargo, en las células deficientes en la ruta FA, la HR no está bien regulada, por lo que la NHEJ se presenta como un mecanismo alternativo de reparación de las DSBs, y éste lleva a las fusiones cromosómicas que permiten la formación de los cromosomas radiales.

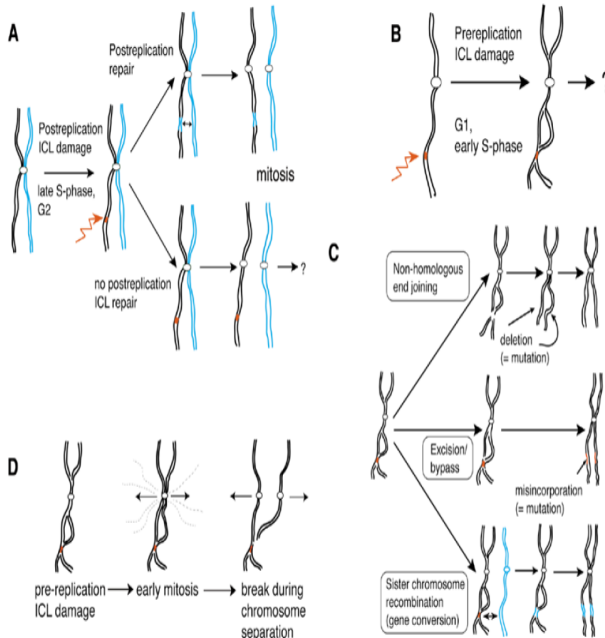


Figura 19. Reparación de ICLs. La figura muestra las diferentes vías de reparación de ICLs durante el ciclo celular. A, ICL producida tras la replicación, que puede ser reparada o no. B, ICL generada antes de la replicación, no hay cromatida hermana disponible. C, mecanismos posibles de reparación de la ICL antes de la replicación. D, una ICL sin reparar genera necesariamente una rotura en las cromatidas durante la mitosis. Adaptado de Grompe & D’Andrea, 2001.

Por otra parte, aunque las ICLs sean reconocidas durante la interfase, se produciría un ciclo fútil de reparación, por el que se generaría una rotura de doble cadena que volvería a ser ensamblada por la maquinaria de la NHEJ, pero sin eliminar la ICL. Debido a todas estas razones, se considera que una de las funciones de las proteínas de la ruta FA/BRCA es precisamente inhibir la vía NHEJ.

Al no existir ninguna proteína que reconozca y elimine específicamente la unión de una cadena de DNA con la otra, la solución para reparar de una manera eficiente las ICLs pasa por coordinar correctamente varias vías de reparación durante la replicación (Niedernhofer *et al.*, 2004; Raschle *et al.*, 2008); el proceso consta de las siguientes etapas, que se suceden de una manera ordenada y coordinada por las proteínas de la vía FA/BRCA (figura 20).

A- Parada de la horquilla de replicación debido a la presencia de la ICL. Es posible que otra horquilla que procede de la dirección opuesta acabe chocando también contra la ICL.

B- FANCD2 recluta las helicasas necesarias y se produce una regresión de la horquilla de replicación, seguida de un corte en una de las cadenas de DNA, generando una DSB. Este corte está mediado seguramente por la acción de la nucleasa FAN1 (Huang *et al.*, 2010).

C-D FANCD2 recluta a SLX4, que media la acción de una endonucleasa, XPF-ERCC1 (o Mus81-Eme1 o SLX1), al otro lado de la ICL (Niedernhofer *et al.*, 2004).

E- Un pequeño fragmento de DNA que contiene la lesión permanece unido a la cadena intacta, mientras que en la cadena complementaria queda un hueco de DNA sin emparejar.

F- FANCD2 recluta una polimerasa de síntesis translesión que sea capaz de pasar a través de la lesión, probablemente para ello llama a la polimerasa primero y luego a PCNA para que ésta pueda polimerizar (Fu *et al.*, 2013).

G- Se recluta la maquinaria de la vía de escisión de nucleótidos (*Nucleotide Excision Repair* o NER), para eliminar el fragmento de DNA que contiene la ICL, dejando expuesto el fragmento de DNA recién sintetizado.

H- La DNA polimerasa ahora es capaz de polimerizar correctamente la cadena complementaria.

I- En la cadena donde se produjo la primera DSB, mientras tanto, se ha producido un procesamiento del extremo, que produce un pequeño fragmento de DNA de cadena sencilla con un extremo 3'-OH libre.

J- El fragmento de ssDNA forma un nucleofilamento con RAD51 y, mediante la acción de BRCA2, invade la cadena complementaria de la doble cadena recién formada.

K- Se restaura la horquilla de replicación y se reinicia la replicación.

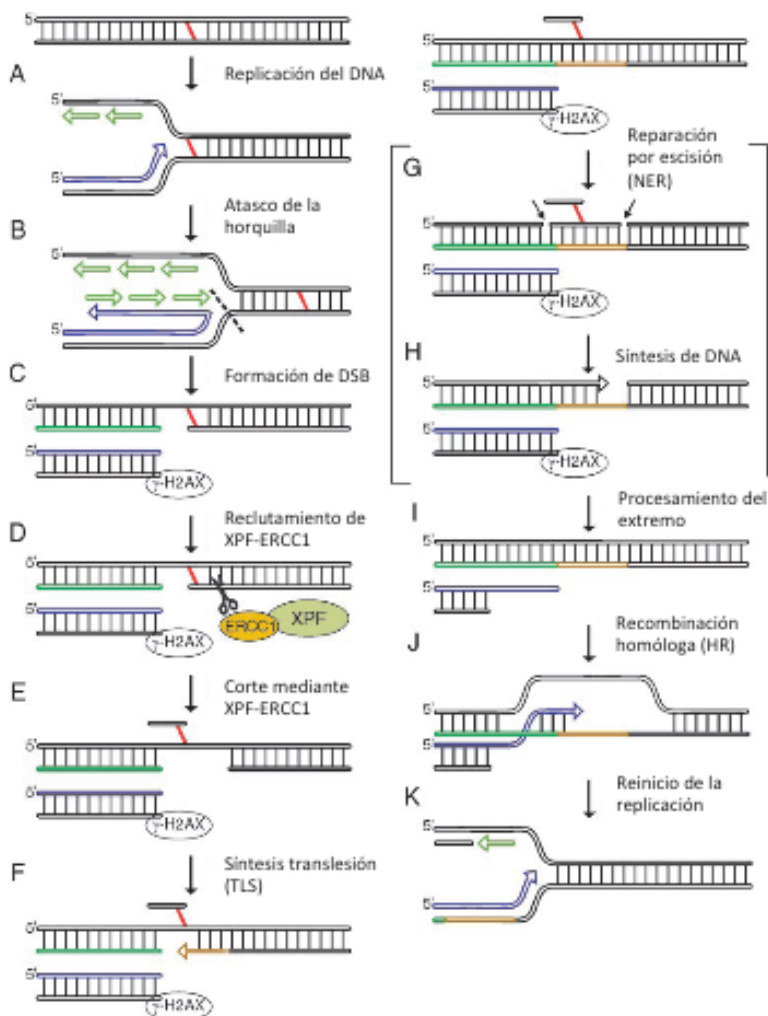


Figura 20. Reparación de entrecruzamientos intercatenar (ICLs) durante la replicación. Detalles en el texto, adaptado de Niedernhofer *et al.*, 2004.

1.3.7.- Anemia de Fanconi y actividad de L1.

La reparación de ICLs es, por tanto, un proceso largo y complejo que se realiza en múltiples etapas, que requiere la participación coordinada de varias vías de reparación, y que genera roturas de doble cadena a modo de lesiones intermedias como parte del proceso de reparación. Asimismo, se ha demostrado que las roturas de doble cadena, que son reparadas por la recombinación homóloga durante el proceso de reparación de ICLs en condiciones normales, no se reparan de forma correcta mediante la HR en las células deficientes en FA, y es la NHEJ la que se encarga de estas lesiones, de una manera que presenta una elevada predisposición a errores en la reparación (Pace *et al.*, 2010).

De hecho, la inhibición de la vía NHEJ en células deficientes en FA es capaz de corregir parcialmente la sensibilidad a mitomicina y el número de aberraciones cromosómicas que se producen, demostrando que la NHEJ corrompe la reparación ante un fallo de la vía FA (Adamo *et al.*, 2010; Pace *et al.*, 2010).

Por otra parte, como se describió en células deficientes en dos componentes principales de la NHEJ, concretamente en XRCC4 (Morrish *et al.*, 2002) y DNA-PKs (Morrish *et al.*, 2007), L1 es capaz de aprovechar las roturas de doble cadena que no se reparan para integrarse cuando la endonucleasa del elemento no es funcional, llegando incluso a presentar niveles de actividad similares a los de elementos con una endonucleasa activa.

Por tanto, cabe pensar que los elementos L1 mutantes en el dominio endonucleasa, de una manera

similar a lo que ocurre cuando falla la vía de reparación NHEJ, son capaces de aprovechar las roturas de doble cadena que no son reparadas de forma eficiente ante fallos de la ruta de la anemia de Fanconi.

Así, en este sistema el elemento encontraría un hábitat favorable para su movilización, lo que podría resultar perjudicial para los pacientes al aumentar su inestabilidad genómica, o por otra parte quizá de forma natural L1 podría estar implicado en la biología de la FA, al aprovechar estas roturas de doble cadena que no son reparadas.

2.- HIPOTESIS Y OBJETIVOS.

2.1.- Hipótesis.

Nuestra hipótesis es que los elementos mutantes en el dominio endonucleasa son capaces de aprovechar las roturas del DNA sin reparar que se producen en las células de los pacientes de Anemia de Fanconi, y que la actividad de los retroelementos se encuentra desregulada en estos pacientes, promoviendo la inestabilidad genómica y aumentando la predisposición al desarrollo de tumores característicos de la enfermedad.

2.2.- Objetivos.

2.2.1.- Determinar si se produce la retrotransposición por la vía ENi en células deficientes en la ruta de la Anemia de Fanconi.

2.2.2.- Caracterizar las inserciones de L1 producidas en células deficientes en la ruta de Anemia de Fanconi.

2.2.3.- Determinar si existe relación entre la actividad de L1 y la inestabilidad genómica presente en la Anemia de Fanconi.

2.2.4.- Analizar los cambios en los niveles de retrotransposición de L1 ante deficiencias en los distintos miembros de la ruta de la Anemia de Fanconi.

MATERIALES Y MÉTODOS

3.- MATERIALES

3.1.- Antibióticos y productos químicos.

Mitomicina C

Es un agente intercalante del DNA, por tanto se deben tomar las precauciones pertinentes durante su manejo. Se obtuvo de Sigma Aldrich (Cat.No. M4287-2MG), procedente de *Streptomyces caespitosus*.

Se prepara una solución *stock*, que tiene color violáceo, con una concentración de 1 mM utilizando 5.98 µg de agua destilada. Posteriormente se realizan diluciones con agua destilada para obtener las concentraciones de 50µM, 100 µM y 200 µM. Para su uso en cultivos celulares se diluye 1000 veces con el medio de cultivo hasta obtener la concentración deseada (50, 100 y 200 nM).

La solución *stock* y las soluciones a distintas concentraciones de uso se mantienen a 4 °C, y tienen una vida media de un mes; pasado ese tiempo la mitomicina pierde efectividad, y se realiza una nueva solución *stock*.

El tratamiento con mitomicina se mantiene durante 72 horas en las células en cultivo, a una concentración final en el medio de 50 nM, 100 nM ó 200 nM.

Blasticidina S HCl

La blasticidina es un antibiótico aminoglucósido, que actúa como inhibidor muy potente de la traducción tanto en procariontas como en eucariotas. Este antibiótico actúa de forma rápida

permitiendo la selección de células resistentes en menos de una semana.

El antibiótico se adquiere de Invitrogen contiene 50 µg, es producido por *Streptomyces griseochromogenes*, número de catálogo R210-01. La solución *stock* se prepara a 100µg/µl, añadiendo 500 µl de agua destilada y filtrando con un filtro de 0.22 µm (Millipore). La solución *stock* se mantiene a -20 °C y es estable durante varios meses.

Para su uso en cultivo, se diluye 1/10000-20000 en medio de cultivo; alternativamente, se puede preparar la solución *stock* a 10 µg/µl, añadiendo 5 ml de agua destilada para 50 µg de Blastidina, y para su uso en cultivo diluyendo 1/1000-2000 en medio de cultivo.

La selección con blastidina se realiza durante 7 días, añadiendo el medio con blastidina tres, cuatro o cinco días tras la transfección (según la naturaleza del experimento). La concentración en medio de cultivo es normalmente de 5 µg/ml en células humanas y 10 µg/ml en células CHO.

Neomicina

Se conoce también como G418 o Geneticina. Su mecanismo de acción es más lento que el de la blastidina, y actúa inhibiendo la síntesis de proteínas mediante su unión al ribosoma, tanto en bacterias como en células eucarióticas. Se obtiene de Gibco (Invitrogen), es producida por *Micromonospora rhodorangea*. La solución *stock* se prepara a 50 µg /ml, y se mantiene a 4 °C, siendo estable hasta 3 meses.

Para la selección de células resistentes se diluye 1/50-250 en medio de cultivo. Para la selección de células HeLa y CHO se utiliza una concentración final de 400 µg/ml (1/125), y para fibroblastos humanos de 200 µg/ml (1/250).

El tratamiento se mantiene durante 12-14 días, reemplazando el medio cada 2 días por medio fresco con Neomicina a las concentraciones especificadas.

Ampicilina

Este antibiótico es un derivado de la penicilina, que causa la parada del crecimiento de las células inhibiendo la síntesis de la pared bacteriana; es por tanto un bacteriostático, ya que no provoca la muerte directa de las bacterias. Se obtiene de Sigma Aldrich, número de catálogo A9518-5G,

La solución *stock* se prepara a 50-100 µg /ml en agua destilada; se pesa 1 g de ampicilina, se diluye en 10-20 ml de agua destilada y se filtra con un filtro de 0.45 µm (Millipore), la solución *stock* se conserva a -20 °C, donde es estable durante varios meses.

Para su uso, tanto en medio líquido como en placa, el medio debe ser primero autoclavado y dejado en reposo a una temperatura inferior a 65 °C (56 °C), para evitar la degradación del antibiótico, tras lo cual se añade el antibiótico diluido 1/500-1000, hasta alcanzar la concentración final de 50-100 µg/ml.

Kanamicina

Este antibiótico, usado para la selección de colonias bacterianas, es el equivalente a la Neomicina utilizada en células mamíferas, con el mismo

mecanismo de actuación. La Kanamicina sulfato utilizada se obtiene de Amresco, número de catálogo 0408-10G. Al contrario que la neomicina, la kanamicina provoca la muerte celular, por lo que es especialmente importante dejar que las células se recuperen tras la transformación para que puedan expresar el gen de resistencia.

La solución *stock* de kanamicina se prepara a 25-50 µg/ml; se pesa 1 g de kanamicina, se diluye en 40-20 ml de agua destilada y se filtra con un filtro de 0.45 µm (Millipore), la solución *stock* se conserva a -20 °C, donde es estable durante meses.

Para su uso, tanto en medio como en placa, el medio LB se autoclava y se deja en reposo a una temperatura inferior a 65 °C (56 °C), para evitar la degradación del antibiótico, tras lo cual el antibiótico se diluye 1/500-1000, alcanzando la concentración final de 25-50 µg/ml.

3.2.- Medios de cultivo bacteriano.

LB (*Luria Broth*)

Adquirido de Sigma, con la siguiente composición:

- 1 l de agua destilada
- 10 g/l de Triptona
- 5 g/l de extracto de levadura
- 5 g/l de NaCl

Preparación:

- Autoclavar el medio.

SOB (*Super Optimal Broth*) con Glicina

Preparado en el laboratorio, con la siguiente composición:

- 1 l de agua destilada

- 20 g/l de Triptona
- 5 g/l de extracto de levadura
- 0.5 g/l de NaCl
- 625µl de KCl 4 M (autoclavado)
- 10 g de Glicina

Preparación:

- Se ajusta a pH 7 con NaOH diluido
- Se autoclava el medio
- Se añaden 10 ml/l de MgCl₂ 1 M (autoclavado)

TB (*Terrific Broth*)

Preparado en el laboratorio, con la siguiente composición:

- 100 ml de agua bidestilada
- 0.30 g de sal de sodio PIPES (ácido 1,4-piperazinodietanosulfónico)
- 0.22 g de CaCl₂*6H₂O
- 1.86 g de KCl

Preparación:

- Los materiales utilizados deben ser enjuagados con agua bidestilada
- Se ajusta el pH a 6.7 usando KOH 10 M
- Se añade 1.09 g de MnCl₂*4H₂O
- Se filtra mediante filtro de 0.45 µm para esterilizar

3.3.- Medios de cultivo celular

DMEM (*Dulbecco Modified Eagle Medium*)

Es un medio para cultivo celular que no contiene proteínas, lípidos o factores de crecimiento, y contiene glucosa. El medio DMEM utilizado contiene rojo fenol como indicador de pH, así como el suplemento Glutamax de Invitrogen; en función de los

requerimientos celulares, se utilizan distintas modalidades:

- Medio DMEM con concentración alta de glucosa (4.5 g/l), que no contiene piruvato, obtenido de Gibco (Invitrogen), número de referencia 10566-016.

- Medio DMEM con una concentración de glucosa baja (1 g/l), que sí contiene piruvato, obtenido de Gibco (Invitrogen), número de referencia 10567-014.

En ambos caso el medio debe ser suplementado, generalmente mediante la adición de suero bovino fetal (FBS), y otros suplementos que varían en función del tipo celular y los experimentos a realizar.

Opti-MEM (*Minimal Essential Media*)

Es un medio mínimo esencial mejorado, que permite el crecimiento celular incluso en ausencia de suero. Este medio está optimizado para su uso conjunto con agentes de transfección lipídicos, obtenido de Gibco (Life technologies), número de referencia 31985-062.

3.4.- Reactivos de transfección lipídicos.

Para los experimentos de transfección de DNA plasmídico, se emplean dos tipos distintos de reactivos dependiendo del tipo celular a transfectar, concretamente **FuGENE 6** de Promega (número de referencia E2691) para las líneas celulares humanas y **X-tremeGENE 9** de Roche (número de referencia 06365787001) para las líneas de Hamster CHO, ambos siguiendo las instrucciones recomendadas por los fabricantes, detalladas más adelante.

Para los experimentos que requieren la transfección de siRNAs se utilizó **DharmaFECT 4**, obtenido de Dharmacon (Thermo), número de referencia T-2004-03, siguiendo las instrucciones recomendadas por el fabricante.

Para los experimentos que requieren la cotransfección de siRNAs y DNA plasmídico se utilizó **DharmaFECTDuo**, obtenido de Dharmacon (Thermo). Este reactivo de transfección lipídico está optimizado para la transfección conjunta de siRNAs y plásmidos en líneas celulares.

3.5.- Soluciones

Solución fijadora formaldehído/glutaraldehído

Preparada en el laboratorio, composición:

- 54 ml de formaldehído 37% (Sigma).
- 4 ml de glutaraldehído 50% (Sigma).
- 100 ml de PBS 10X (Gibco).
- 842 ml de agua destilada.

Solución fijadora ácido acético/metanol

Preparada en el laboratorio, composición:

- 100 ml de ácido acético.
- 100 ml de metanol.
- 800 ml de agua destilada.

Solución de cristal violeta

Esta solución permite la tinción del núcleo de las células de mamíferos sin necesidad de permeabilización previa. El cristal violeta es un agente intercalante del DNA, y por lo tanto debe ser utilizado con las precauciones correspondientes. Su

composición es de 0.1% peso/volumen de cristal violeta en agua; es decir, 1 g de cristal violeta en 1 l de agua destilada. Alternativamente, se puede preparar una solución de cristal violeta 1% peso/volumen de cristal violeta en metanol (1 g de cristal violeta en 100 ml de metanol).

Solución de disolución

Esta solución permite disolver el cristal violeta de células previamente fijadas y teñidas, debe emplearse únicamente cuando las placas que van a ser tratadas están secas. Su composición es de 0.1% peso/volumen de SDS en metanol.

3.6.- Cepas bacterianas.

DH5-alfa

Es la cepa de *E.coli* más utilizada para la transformación y amplificación de vectores, válida tanto para electroporación como para la transformación por choque térmico tras su crecimiento en las condiciones apropiadas (ver más abajo), obtenidas del Banco Andaluz de Células Madre.

XL10-GOLD

Obtenida de Stratagene, esta cepa de *E.coli* está optimizada para presentar una alta eficiencia de transformación, con más de 1×10^9 unidades formadoras de colonias por microgramo de DNA. Permite la transformación química por choque térmico de vectores de un tamaño considerable (de

más de 10 kb), y está diseñada para permitir la amplificación (replicación) de DNA humano metilado.

3.7.- Líneas celulares

3.7.1.- Humanas.

HELA

Células de carcinoma de cuello de útero, es la primera línea celular que se consiguió aislar y mantener en cultivo (Bang et al., 1951) y la más utilizada, a pesar de tener un cariotipo alterado muy diferente a las células normales. Cedidas por el laboratorio del Dr. John Moran (University of Michigan, EEUU).

HEK293T

Línea celular embrionaria de riñón, extremadamente eficiente en la producción de partículas virales. Cedidas por el laboratorio del Dr. Pablo Menendez (Genyo, España).

U2OS

Células de osteosarcoma, que destacan por presentar un citoplasma de gran tamaño, lo que las hace muy útiles para estudios de localización celular. Cedidas por el laboratorio del Dr. John Moran (University of Michigan, EEUU).

PD20F

Fibroblastos obtenidos de un paciente de Anemia de Fanconi, y pertenecientes al grupo de complementación D2, pueden ser obtenidas del repositorio del instituto Coriell, número de catálogo

GM16633. Cedidas por el laboratorio del Dr. Niall Howlett (University of Rhode Island, EEUU).

PD20F + FANCD2

Obtenidas a partir de las PD20F tras complementación con el gen FANCD2 mediante un vector retroviral; esta línea contiene el gen FANCD2 humano, que corrige el defecto presente en la línea parental. Cedidas por el laboratorio del Dr. Niall Howlett (University of Rhode Island, EEUU).

ANKN

Fibroblastos obtenidos de un paciente de Anemia de Fanconi, cedidos por el laboratorio de Arleene Auerbach, pertenecientes al grupo de complementación I. Cedidas por el laboratorio del Dr. Niall Howlett (University of Rhode Island, EEUU).

ANKN + FANCI

Obtenida a partir de la línea ANKN, contiene el gen FANCI introducido mediante el uso de un vector retroviral, que corrige el defecto presente en la línea parental. Cedidas por el laboratorio del Dr. Niall Howlett (University of Rhode Island, EEUU).

ANKN + LACZ

Obtenida a partir de la línea ANKN, contiene el gen lacZ introducido mediante un vector retroviral y posee el defecto presente en la línea parental. Cedidas por el laboratorio del Dr. Niall Howlett (University of Rhode Island, EEUU).

VU423

Fibroblastos obtenidos de un paciente de Anemia de Fanconi, pertenecientes al grupo de complementación BRCA2. Cedidas por el laboratorio del Dr. Niall Howlett (University of Rhode Island, EEUU).

GM16635

Fibroblastos obtenidos de un paciente de Anemia de Fanconi, pertenecientes al grupo de complementación FANCG, obtenidas del repositorio del instituto Coriell, número de catálogo GM16635. Cedidas por el laboratorio del Dr. Niall Howlett (University of Rhode Island, EEUU).

HCT-116

Línea tumoral epitelial humana de cáncer colorrectal, obtenida del repositorio de la ATCC (*American Type Culture Collection*), número de catálogo CCL-247. Cedidas por el laboratorio del Dr. Bert Vogelstein (John Hopkins Med School, EEUU).

3.7.2.- Hámster

Las líneas empleadas en este estudio son derivadas de una línea de fibroblastos de pulmón de hamster (V79 y V79B), a excepción de las líneas KO40 y 40B6 (Tebbs et al., 2005), que fueron derivadas de una línea CHO (por *Chinese Hamster Ovary*). La peculiaridad más notable de las células derivadas de CHO es que poseen un genoma pseudodiploide, lo que las hace que la mutación de una sola copia de un gen puede producir un fenotipo. Han sido especialmente relevantes en el campo de la reparación de DNA,

debido a la relativa facilidad para producir mutantes y una elevada viabilidad, permitiendo realizar ensayos de *screening* rápidos mediante experimentos sencillos; además, muchas líneas celulares diploides presentan problemas muy notables en su proliferación ante una deficiencia grave en algunos de estos genes. Por otro lado, aquellas líneas derivadas de pulmón de Hamster son diploides al ser derivadas de fibroblastos de pulmón.

La mayoría de las células CHO derivan de la línea CHO-K1, aislada por primera vez en 1958 (Puck, 1958) y cuyo genoma se terminó de secuenciar en el año 2011 (Xu et al., 2011), aunque no existe una versión completa de su ensamblaje finalizada.

V79

Derivada originalmente de pulmón de Hámster, de fenotipo normal, pueden ser obtenidas del repositorio del instituto Coriell, número de catálogo GM00215. Cedidas por el laboratorio de la Dra. Margaret Zdzienicka (Leiden University, Países Bajos).

V79B

Derivada originalmente de pulmón de Hámster, es una sublínea clonal diferente de las V79, de fenotipo normal, pueden ser obtenidas del repositorio del instituto Coriell, número de catálogo GM16136. Cedidas por el laboratorio de la Dra. Margaret Zdzienicka (Leiden University, Países Bajos).

VC8

Derivada de la línea V79B, esta línea posee fenotipo FA, que puede ser corregido por la expresión del gen

BRCA2. Cedidas por el laboratorio de la Dra. Margaret Zdzienicka (Leiden University, Países Bajos).

VH4

Derivada de la línea V79B, esta línea posee fenotipo FA, que puede ser corregido parcialmente mediante la expresión del gen humano FANCA. Cedidas por el laboratorio de la Dra. Margaret Zdzienicka (Leiden University, Países Bajos).

VH4 + FANCA

Obtenida a partir de la línea VH4 tras infección con un retrovirus que contiene el gen FANCA humano, que corrige la deficiencia presente en la línea parental.

CLV5B

Derivada de la línea V79, esta línea celular posee fenotipo FA, que puede ser corregido mediante la expresión del gen humano FANCC. Cedidas por el laboratorio de la Dra. Margaret Zdzienicka (Leiden University, Países Bajos).

CLV5B + FANCC

Obtenida a partir de la línea CLV5B, tras infección con un retrovirus que contiene el gen FANCC humano, que corrige la deficiencia presente en la línea parental.

40B6

Derivada de la línea CHO-K1, de fenotipo normal. Cedida por el Dr. Larry Thompson (Lawrence Livermore National Laboratory, EEUU).

K040

Derivada de la línea 40B6, esta línea celular posee fenotipo FA, que puede ser corregido mediante la expresión del gen humano FANCG. Cedida por el Dr. Larry Thompson (Lawrence Livermore National Laboratory, EEUU).

3.8.- Vectores y plásmidos recombinantes.

Todos los vectores utilizados fueron derivados del pCEP4 de Invitrogen, a no ser que se indique lo contrario. El **pCEP4** es un vector de expresión eucariota que contiene el promotor del citomegalovirus (CMV) aguas arriba del sitio de clonaje y una señal de poliadenilación del virus SV40 aguas abajo del sitio de clonaje. Este vector también contiene un origen de replicación eucariota, necesario para mantenerse episomalmente en las células transfectadas.

El elemento LINE-1 utilizado en este estudio fue el elemento L1.3, y los siguientes mutantes de éste:

- **L1.3 D205A**, con una mutación puntual en el dominio EN de ORF2 que destruye la actividad endonucleasa pero no afecta a la estabilidad de la ribonucleopartícula (RNP).
- **L1.3 H230A**, con una mutación puntual en el dominio EN de ORF2 que destruye la actividad endonucleasa; algo más severa que la mutación D205A, posiblemente por un efecto sobre la estabilidad de la RNP.
- **L1.3 D702A**, con una mutación puntual en el dominio RT de ORF2 que destruye la actividad

reversotranscriptasa pero no afecta a la estabilidad de la ribonucleopartícula (RNP).

Los vectores utilizados fueron los siguientes:

- TAM102/L1.3

Estos vectores contienen la secuencia completa del ORF1, ORF2 y 3'UTR del elemento L1.3, salvaje o mutante según indicado, sin la secuencia 5'UTR, y un casete de resistencia a blasticidina aguas abajo del elemento (Morrish et al., 2002).

- JM101/L1.3

Este vector contiene la secuencia completa del L1.3 (incluida la 5'UTR) y un cassette de resistencia a neomicina aguas abajo del elemento (Morrish et al., 2002).

- JJ101/L1.3

Estos vectores, a diferencia de los TAM102, contienen la secuencia completa del L1.3 incluida la 5'UTR, salvaje o mutante según indicado, y un casete de resistencia a blasticidina aguas abajo del elemento (Beck et al., 2010).

- TAM102/L1.3 k7i

Estos vectores contienen la secuencia completa del ORF1, ORF2 y 3'UTR del elemento L1.3, salvaje o mutante en el dominio endonucleasa según indicado, sin la secuencia 5'UTR, y el casete de recuperación de inserciones (k7i), compuesto por un gen de resistencia a neomicina en células de mamíferos y kanamicina en células bacterianas interrumpido por un intrón eucariótico, y aguas abajo un origen de replicación bacteriano que permite la replicación en bacterias.

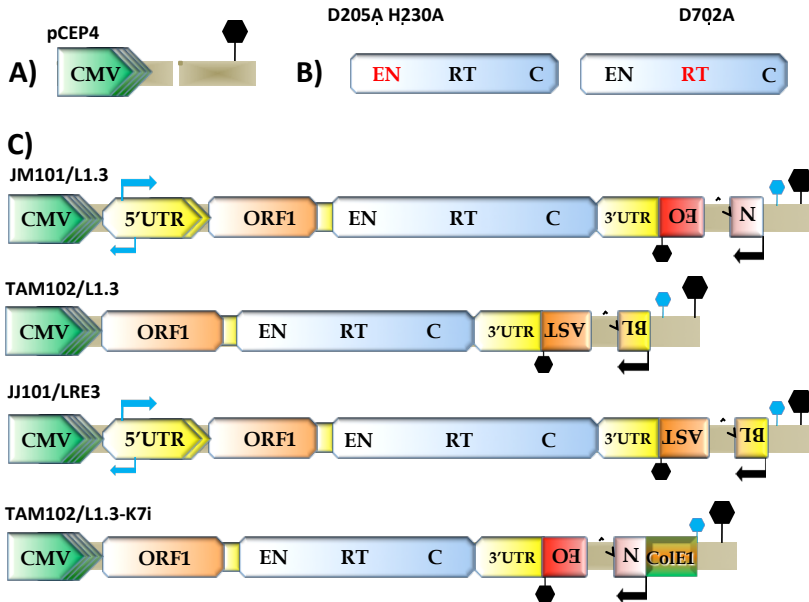


Figura 21. Vectores empleados en este estudio.

A) Sitio de clonaje del vector pCEP4. CMV indica el promotor del citomegalovirus, el hexágono negro indica la señal de poliadenilación del virus SV40. **B)** Posiciones relativas de las mutaciones de L1 en el ORF2, se indica en rojo el dominio afectado por la mutación; “EN” indica endonucleasa, “RT” indica reversotranscriptasa y “C” dominio rico en cisteínas. **C)** Esquema de los elementos presentes en los diferentes vectores de expresión de L1 empleados; las flechas azules indican la posición de las regiones promotoras de la 5’UTR y su dirección y fuerza relativas, el hexágono azul indica la señal de poliadenilación de L1, en rojo “NEO” indica el gen de resistencia a neomicina, en naranja “BLAST” indica el gen de resistencia a blasticidina, en verde y naranja “ColE1” indica origen de replicación bacteriano.

- pCEP4-eGFP

Este vector contiene el gen de la proteína verde fluorescente sin interrumpir; se utiliza en cotransfección como control, para medir la eficiencia de transfección del plásmido co-transfectado.

- pcDNA6/myc-his

Este vector no procede del pCEP4 y no se mantiene episomalmente en las células transfectadas; contiene el gen de resistencia a blasticidina sin interrumpir, con su promotor y señal de poliadenilación.

- pBABE/Hygro-hTERT

Este vector no procede del pCEP4 y no se mantiene episomalmente en las células transfectadas; contiene el gen de la telomerasa humana sin interrumpir bajo el control de un promotor CMV, su propia señal de poliadenilación, y el gen de resistencia a higromicina; ambos genes se encuentran flanqueados por las LTRs necesarias para la producción de partículas retrovirales.

3.9.- RNAs de interferencia.

Obtenidos de Dharmacon (Thermo), con el nombre comercial “ON-TARGET plus SMART pool”, consisten en un pool de 4 siRNAs diferentes para un mismo RNA diana, lo que mejora la eficacia en la disminución específica (*knockdown*) de la expresión de dicho RNA. En la figura 22 se muestra un esquema de las proteínas implicadas en reparación de DNA estudiadas así como los siRNAs empleados y el proceso de reparación afectado principalmente por la inhibición de cada componente.

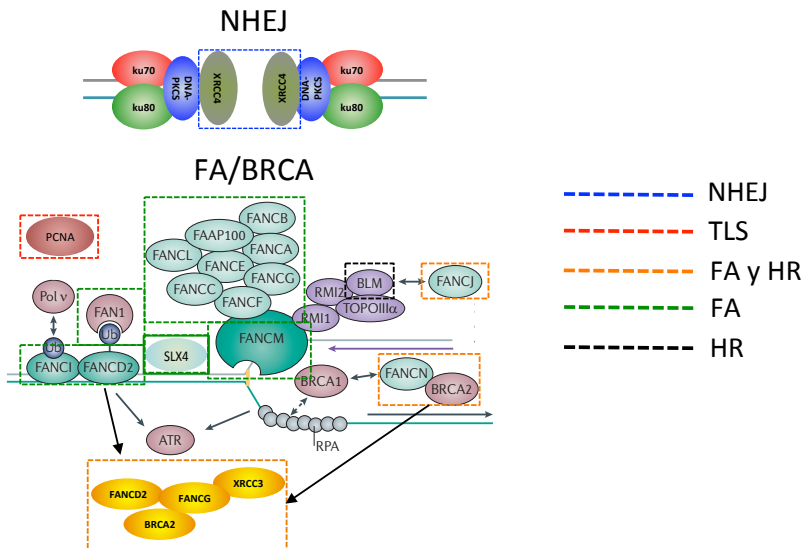


Figura 22. Esquema de los procesos de reparación del DNA inhibidos por los siRNAs empleados en el estudio. La figura muestra un esquema general de las rutas de reparación y proteínas analizadas, indicando mediante un recuadro de color el proceso afectado principalmente por la inhibición de la proteína. Abreviaturas: NHEJ, Unión de extremos no homólogos; TLS, síntesis translesión; FA (Anemia de Fanconi), reconocimiento y coordinación del procesamiento de la lesión (incluye TLS); HR, recombinación homóloga.

Concretamente, se utilizaron los siguientes siRNAs:

- Control siRNA no diana (***Non-targeting*** siRNA), número de referencia D-001810-10-20, ON-TARGETplus Non-targeting Pool.
- Control siRNA frente a **XRCC4**, número de referencia L-004494-00-0010, ON-TARGETplus SMARTpool, Human XRCC4 (7518).
- siRNA frente a **FANCA**, número de referencia L-019283-00-0005, ON-TARGETplus SMARTpool, Human FANCA (2175).
- siRNA frente a **FANCB**, número de referencia L-016941-00-0005, ON-TARGETplus SMARTpool, Human FANCB (2187).
- siRNA frente a **FANCC**, número de referencia L-011033-00-0005, ON-TARGETplus SMARTpool, Human FANCC (2176).
- siRNA frente a **FANCD1/BRCA2**, número de referencia L-003462-00-0010, ON-TARGETplus SMARTpool, Human BRCA2 (675).
- siRNA frente a **FANCD2**, número de referencia L-016376-00-0010, ON-TARGETplus SMARTpool, Human FANCD2 (2177).
- siRNA frente a **FANCE**, número de referencia L-013991-00-0005, ON-TARGETplus SMARTpool, Human FANCE (2178).
- siRNA frente a **FANCF**, número de referencia L-014206-00-0005, ON-TARGETplus SMARTpool, Human FANCF (2188).
- siRNA frente a **FANCG**, número de referencia L-011899-00-0005, ON-TARGETplus SMARTpool, Human FANCG (2189).

- siRNA frente a **FANCI/KIAA1794**, número de referencia L-022320-01-0005, ON-TARGETplus SMARTpool, Human KIAA1794 (55215).
- siRNA frente a **FANCI/BRIP1**, número de referencia L-010587-00-0005, ON-TARGETplus SMARTpool, Human BRIP1 (83990).
- siRNA frente a **FANCL**, número de referencia L-021486-00-0005, ON-TARGETplus SMARTpool, Human FANCL (55120).
- siRNA frente a **FANCM**, número de referencia L-021955-00-0005, ON-TARGETplus SMARTpool, Human FANCM (57697).
- siRNA frente a **FANCN/PALB2/FLJ21816**, número de referencia L-012928-01-0005, ON-TARGETplus SMARTpool, Human FLJ21816 (79728).
- siRNA frente a **FANCO/RAD51C**, número de referencia L-010534-00-0005, ON-TARGETplus SMARTpool, Human RAD51C (5889).
- siRNA frente a **SLX4**, número de referencia L-014895-00-0005, ON-TARGETplus SMARTpool, Human BTBD12 (84464).
- siRNA frente a **FAN1**, número de referencia L-020327-00-0005, ON-TARGETplus SMARTpool, Human KIAA1018 (22909).
- siRNA frente a **BLM**, número de referencia L-007287-00-0005, ON-TARGETplus SMARTpool, Human BLM (641).
- siRNA frente a **PCNA**, número de referencia D-003289-05-0005, ON-TARGETplus SMARTpool, Human PCNA (5111).

3.10.- Oligonucleótidos sintéticos.

Para la correcta caracterización de las inserciones de L1, descrita en métodos, se precisa la realización de varias rondas de secuenciación capilar. En la figura 23 se muestra un esquema de la posición relativa de los oligonucleótidos sintéticos o *primers*, utilizados para la secuenciación y el subsiguiente análisis de las inserciones.

Oligonucleótidos situados en el casete de recuperación de inserciones y la secuencia de L1:

- **Reco3** se sitúa en el origen de replicación bacteriano y se dirige hacia la cola de poliA creada en el evento de retrotransposición.
- **Neo210as** se sitúa en el gen de resistencia a neomicina / kanamicina y se dirige aguas arriba de la inserción, es decir, hacia la secuencia del elemento L1.
- **JORF2"X"as**: están situados en la secuencia del ORF2 del L1 en orden ascendente, esto es, la letra "A" marca el que está más cerca del ORF1 de L1 y la letra "J" el que está más cerca del gen neo, "as" indica que la dirección del primer es hacia aguas arriba. En la secuenciación se utilizan primero los más cercanos al gen neo hasta encontrar la región genómica en la que se produjo la inserción, así como el punto de truncación de la secuencia de L1.
- **JORF2"X"sense**: están situados en la secuencia del ORF2 del L1 en orden ascendente también, en la misma posición que los *primers* "as" pero en dirección opuesta, "sense" indica que la dirección es hacia aguas abajo. Se utilizan cuando se encuentra una inversión/truncación de L1 originada por un evento

de *twin priming*, para poder alcanzar la región genómica donde se produjo la inserción.

- *Primers* situados en la región genómica donde se produjo la inserción contenidos en el vector recuperado.

En muchas ocasiones la secuencia obtenida de la secuenciación con el primer Reco3 es de buena calidad hasta alcanzar la cola de poliA, pero la presencia de una gran cantidad de Adeninas suele provocar que la Taq se deslice y por tanto la calidad de la secuencia aguas abajo es demasiado baja, haciendo imposible conseguir un alineamiento fiable de la secuencia genómica situada en 3'.

Cuando esto ocurre, y una vez se ha alcanzado la región genómica contigua a la inserción en 5', es necesario el uso de *primers* para alcanzar la cola de poliA desde el otro extremo del plásmido recuperado. Llamamos a estos *primers* como "*Walking*" o "*W*", indicando en el primer número el clon recuperado (o vector que contiene una inserción recuperada), "*Wx*" indica que la secuencia pertenece a la localización genómica en la que se produjo la inserción, contenida en el clon recuperado, "*x*" indica el número de primer distinto situado en la misma región genómica. Se indican los *primers* en orden según el número de clon en el que se utilizaron.

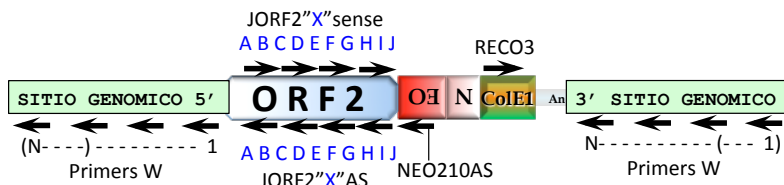


Figura 23. Esquema de los oligonucleótidos empleados en la caracterización de inserciones. Se indica la posición relativa de los oligonucleótidos (*primers*) utilizados para la secuenciación de las inserciones (más detalles en métodos).

Primers en la inserción: a continuación se indica la secuencia nucleotídica de cada uno de los oligonucleótidos sintéticos utilizados en este estudio para la secuenciación de las inserciones recuperadas de L1.

RECO3: CTAAAGTATATATGAGTAACC
 NEO210as: GACCGCTTCCTCGTGCTTTACG
 JORF2]sense: CATTAGTGGGTGCAGCGCACC
 JORF2]as: GGTGCGCTGCACCCACTAATG
 JORF2]sense: GACTGGATTAAGAAAATGTG
 JORF2]as: CACATTTTCTTAATCCAGTC
 JORF2Hsense:GATATCATCTCACACCAGTTAG
 JORF2H]as: CTAACTGGTGTGAGATGATATC
 JORF2G]sense:CACCTATACAAAAATCAATTC
 JORF2G]as: GAATTGATTTTTTGTATAAGGTG
 JORF2F]sense:GGAAGAATCAATATCGTGAAAATG
 JORF2F]as: CATTTTCACGATATTGATTCTTCC
 JORF2E]sense:CAATAAATTAGGTATTGATGGGAC
 JORF2E]as: GTCCCATCAATACCTAATTTATTG
 JORF2D]sense:CCAGATGGATTCACAGCCGAATTC
 JORF2D]as: GAATTGCGCTGTGAATCCATCTGG
 JORF2C]sense:GACACCCTAACATCACAATTA

JORF2Cas: TTAATTGTGATGTTAGGGTGTG
ORF2Bsense: CAGAACTCTCCACCCCAAATC
ORF2Bas: GATTTGGGGTGGAGAGTTCTG
ORF2Asense: ATCAAAAGAGACAAAGAAGGC
ORF2Aas: GCCTTCTTTGTCTCTTTTGAT

Primers dirigidos hacia el sitio genómico: cada uno es específico para una región concreta del genoma (humano o de hámster). El primer número indica el número de clon o vector que contiene una inserción recuperada, “Wx” indica que el primer se ubica en una región genómica, donde “x” indica el número de primer distinto situado en la misma región genómica. Se indican los *primers* en orden según el número de clon en el que se utilizaron (más detalles en métodos).

1-W1: GTCATCAAAAAAGCTCTCCTGC
3-W1: GCAGCTCATTCTAAGAATC
5-W1: GCTTCTCTGCTCAGAACCCTGC
7-W1: GCTGGTTCAGAAAGCATTTTGG
7-W2: GCTTTGGCTGGGGCCAGGC
7-W4: GCAACATGTAAGTAGAGTC
19-W1: GAGGTAGAATTAGCCTCTCC
19-W2: GGAGGGTGGCTAGGAAAGGC
19-W3: GGGAGGGTGGCTAGGAAAGGC
21-W1: GCCTCTCCCATATCCTTCCCTG
22-W1: CCATTGTCATTTCAATTTGATATAC
24-W1: CTGAGAAAGGGTTGAACATC
30-W1: GCCTGTAGGGCTCTTTGTCAAC
31-W1: CTTATCACATGTAGCAGTAAC
31-W2: GCTGGTGTGAACTCCTGGGC
31-W3: GGTACTCTGTTAGCCCCTGGC
31-W4: GCTGGTGTGAACTCCTGGGC

33-W1: CAGCTTGATGATAAGCTATG
34-W1: GACCTACTCGATTTCTTCGTTT
35-W1: GATGATCACCATTCTAACTGG
39-W1: GAGAGATTTAATGATTTTATG
40-W1: CAAGCAGGATGCCATTCTCAC
41-W1: CAGTTTTCTGTTATTTACAGC
44-W1: CAAATGCTCTCGATAATGAG
54-W1: CCGTGTGCAGCAGCTGGTGC
54-W2: CTCAAGACCTGTGCAAGAGC
54-W2B: CAAGAGTGCTAGTTGTAAGC
54-W3: CAGAAGTTAGTAGAGAGGTGC
55-W1: GTGAAGGAAGTACTGGAAACAG
55-W2: CTGTTTCCAGTACTTCCTTCAC
55-W3: CAGTTTGAGTGTCCCTGTTTTGC
55-W4: GCTTTAGTAATGTGGGTGCC
55-W5: GTCTATTTGGTGTCTGAAAGC
57-W1: GTAATAACTAGTGTACTAGC
57-W2: GCACGGGCCAGACAAGTTGGC
58-W1: GTCAAAGGATAACCCTATGTGC
60-W1: CTATGAGAACAGTGTCCCGTG
64-W1: CATATCCATCTGCTGCAAAGC
64-W2: CAAGCCTACAAAGAATGCATG
65-W1: GCATTATGCCTTAGAGGAGC
65-W2: CAATGGAACAGCATTGCTGGC
65-W3: CATGTGTTATCTACGTGCCG
65-W4: GAAATCCTCAGTCTGTGCCG
71-W1: GCTGGATAAGCTGTTGTTTACG
76-W1: GTACACCACCACATCTGGGC
76-W2: GTCTAGTTGACAAATAAGGC
76-W3: CTAAGATTTGACAGGATTGGC
76-W4: CCTGCCAGCCCCAGGTAAAGC
80-W1: CTTGTATTTCTGGTCTAGGGC
80-W2: GCAAACAGCCATCTAGCTGC
84-W1: GCAAGTATAAAGGTAAGGATTTG

84-W2: CAATGATTACCTAATATTAGC
85-W1: CGGAGTTTTGCTCTGTCGCC
90-W1: AACCAATTTCTCTTCCCTAGC
94-W1: ATTTAACTTTGACTTCTTGC
94-W2: GCCTTGTTCCTACAAATGC
94-W3: GCTAAAGACTTATGCAGAGC
96-W1: CACAGGAAATAGATGGATAGC
96-W2: GTCCTGCCGCCGGCCAGGC
96-W3: GGGTGCCAGAGAAAACCTGGC
97-W2: CTACAAATTGCCCAAGGCGGC
97-W3: GTCCAGCTCCCAGCCACGTGC
97-W4: GACCTAGTTTTGTGTCCCAGC
100-W1: CTGGACCTGGAATTTGCAGC
101-W1: GCACAGAATTTATCAGAGTAAC
106-W1: GCAAACATTGGTATATTGGC
108-W1: GCTAGTCTTTAAGCTCTGAC
109-W1: GCTTCACCTTTTAATACATG
114-W1: GCTAGATGCCCACTGGAATC
115-W1: GTAATCAATCATTGATAGACTG
117-W1: CTCTTCTTGTGCTATCACAGC
134-W1: GTACTAAAGTACTTCCAGCC
149-W1: GATATATGAGCAACAATTACTG

4.- MÉTODOS

4.1.- cultivos celulares.

Células Hela, U2OS y HEK293T

Estas líneas celulares transformadas se cultivan en medio DMEM sin piruvato con 4.5 g/l de glucosa, suplementado con Glutamax, 10% de Suero Bovino Fetal (*Fetal Bovine Serum*, FBS) y Penicilina/Estreptomicina (*Penicillin/Streptomycin*, P/S). Aproximadamente una vez por semana mediante tripsinización se expanden en una proporción de entre 1/10-1/40, típicamente cuando la confluencia se acerca al 100%, y buscando una confluencia no inferior al 5-10% para que sigan creciendo en fase logarítmica.

Reactivo de transfección de DNA: FuGENE 6, normalmente se obtiene una eficiencia de transfección entre 40-80% dependiendo de la línea celular y el tamaño del plásmido, utilizando 1 µg de DNA por pocillo en placa de 6 pocillos.

- Selección con neomicina (G418): ≥ 400 µg/ml
- Selección con blasticidina: 10 µg/ml

Fibroblastos primarios de Anemia de Fanconi

Se cultivan en medio DMEM sin piruvato con 4.5 g/l de glucosa, suplementado con Glutamax, 20% de FBS y P/S, se expanden aproximadamente una vez cada 15-20 días mediante tripsinización, en una proporción 1/2-1/6, necesariamente cuando la confluencia está en torno al 100%, buscando una confluencia en torno al 30-50%. La inmortalización debe realizarse durante los primeros pases, ya que la

proporción de células senescentes aumenta con pases sucesivos.

- Selección con higromicina: 100 µg/ml

Fibroblastos inmortalizados de Anemia de Fanconi

Se cultivan en medio DMEM sin piruvato 4.5 g/l de glucosa, suplementado con Glutamax, 10-20% de FBS y P/S, se expanden aproximadamente una vez cada 10-15 días mediante tripsinización en una proporción 1/5-1/10, habitualmente cuando la confluencia está en torno al 100%, buscando una confluencia en torno al 20-40%.

Reactivo de transfección de DNA: FuGENE 6 o X-tremeGENE 9. La eficiencia de transfección es relativamente baja, generalmente en torno al 10-20%, utilizando 1 µg de DNA por pocillo en placa de 6 pocillos.

Selección con neomicina (G418): 200 µg/ml

Selección con blasticidina: 5 µg/ml

Selección con higromicina: 100 µg/ml

4.2.- Preparación de bacterias competentes.

Se refresca la cepa bacteriana sembrando en estrías en una placa de LB sin antibiótico, tras lo cual se eligen 2 o 3 colonias y se ponen a crecer en agitación, 16 horas a 37 °C en 2 o 3 tubos de polipropileno (cada colonia en un tubo independiente) con 4 ml de medio LB sin antibiótico; posteriormente, se traspasan los 4 ml del cultivo a un Erlenmeyer de 1 l de capacidad con 250 ml de medio

SOB suplementado con glicina. Para su correcto crecimiento, las bacterias se cultivan a 18 °C con agitación suave, hasta que la densidad óptica OD₆₀₀ alcanza un valor de de aproximadamente 0.94, normalmente esto ocurre tras 18-24h.

Una vez alcanzada la densidad óptica óptima, se puede continuar con la preparación, para lo cual es crítico mantener las bacterias a 4 °C o una temperatura inferior en todo momento, trabajando en cámara fría si es preciso.

Primero se incuban las bacterias en hielo durante 10 minutos, tras lo cual se pasan a tubos de 50 ml y se centrifugan a 2500 x g durante 10 minutos a 4 °C. Al terminar la centrifugación se decanta el sobrenadante y se resuspenden las células con suavidad en 80ml de medio TB (*Terrific Broth*) frío mediante repipeteo, y se dejan en hielo 10 minutos.

Posteriormente se repite la centrifugación anterior, a 2500 x g durante 10 minutos a 4 °C, y se decanta el sobrenadante de nuevo, pero tras esta segunda centrifugación las células se resuspenden en 10 ml de medio TB frío. Tras resuspender las células se añaden 10 ml más de medio TB frío que contiene un 14% (v/v) de DMSO y se dejan en hielo otros 10 minutos más.

Pasado este tiempo, se invierte el tubo 2 o 3 veces con suavidad para obtener una suspensión homogénea y se procede a alicuotar las bacterias en tubos de 1.5 ml colocados en etanol con hielo seco (alternativamente se puede utilizar nitrógeno líquido). Por último se guardan a -80 °C.

4.3.- Transformación de plásmidos superenrollados.

La transformación de las bacterias competentes con el vector correspondiente se realiza mediante choque térmico, brevemente:

- Se descongela un vial de bacterias competentes en hielo.
- Se colocan 50-100 ng de plásmido (normalmente 1-5 μ l) en un tubo de 1.5 ml con 10-50 μ l de bacterias competentes y se deja en hielo durante 30 minutos.
- Se realiza el choque térmico, que consiste en colocar el tubo con las bacterias y el vector a 42 °C durante 45 segundos en el caso de las bacterias DH5- α (el tiempo de choque térmico varía para otras cepas), típicamente el proceso se realiza en un baño de agua.
- Inmediatamente tras el choque se coloca el tubo en hielo durante 2 minutos, para recuperar las bacterias del choque térmico.
- Se añaden 100-500 μ l de medio LB precalentado a 37 °C sin antibiótico y se deja recuperar las bacterias con o sin agitación durante 1 hora a temperatura ambiente o 37 °C, para permitir la expresión de los genes de resistencia.
- Se coloca una cantidad apropiada de bacterias en una placa con el antibiótico apropiado; típicamente se plaquea entre una quinta y una décima parte del total de la transformación (10-50 μ l), y se esparce homogéneamente por la placa mediante una espátula de Driglasky.
- Se incuba la placa a 37 °C 12-16h . Las colonias resultantes se utilizan para el aislamiento de DNA plasmídico.

4.4.- Aislamiento y purificación de DNA plasmídico.

El aislamiento de DNA plasmídico de bacterias es necesario para la obtención de los plásmidos utilizados en las transfecciones, así como para la obtención de plásmidos recombinantes del sistema de recuperación de inserciones.

Para la obtención de cantidades moderadamente bajas de plásmido, se realiza un proceso genéricamente llamado *miniprep*, que permite la obtención de cantidades de DNA a una concentración de aproximadamente 50-200 ng/μl. En este estudio se empleó un kit de extracción comercial de Promega, número de referencia A1460, siguiendo las indicaciones del fabricante.

Para la obtención de concentraciones relativamente altas de plásmido, el proceso se conoce como *midiprep*, y permite obtener concentraciones entre 250-1000 ng/μl de DNA plasmídico superenrollado; para ello utilizamos un kit de extracción de Qiagen, número de referencia 12143.

Brevemente, este proceso se basa en la lisis alcalina de las bacterias que contienen el plásmido, neutralización y el posterior uso de unas columnas que contienen una resina a las que el DNA se une de manera específica. Tras su lavado, se procede a la elución del DNA mediante el uso de agua destilada en el caso de las minipreps o de un buffer de elución específico en el caso de las midipreps.

Para la extracciones de DNA plasmídico, se selecciona una colonia aislada y se inoculan las bacterias en una cantidad apropiada de medio LB con

el antibiótico de selección específico, a una concentración apropiada (100 µg/ml de Ampicilina o 25 µg/ml de kanamicina).

Para el crecimiento de las bacterias en el medio líquido se emplearon:

- Para la realización de minipreps: 2 ml de LB con antibiótico, en tubos de polipropileno de 14 ml, con tapones de polietileno que permiten 2 posiciones (para aerobiosis o anaerobiosis), obtenidos de VWR, número de referencia 60818-725. Estos tubos son esteriles.

- Para la realización de midipreps: 50 ml de LB con antibiótico, en frascos Erlenmeyer de vidrio de capacidad 250 ml, obtenidos de Pyrex, número de referencia 29174-149 (VWR). Se utiliza papel de aluminio a modo de tapón para evitar la contaminación y permitir la aireación del cultivo. Los frascos se esterilizan mediante autoclave.

4.5.- Aislamiento y purificación de DNA genómico.

El aislamiento de DNA genómico es necesario para obtener la localización y características de las inserciones de L1, recuperadas mediante el sistema de recuperación de inserciones o *recovery*.

Brevemente, este proceso consiste en el aislamiento de los núcleos celulares, tras lo cual se procede a una lisis de los mismos, y se usan unas columnas a las que el DNA de alto peso molecular (20-150 kb) se une de manera específica, para ello, utilizamos el kit "*Blood and Cell Culture DNA midi kit*" de Qiagen, número de referencia 13343, siguiendo las

instrucciones del fabricante. Tras la elución con 5 ml del buffer correspondiente (QF), realizamos la precipitación del DNA mediante la adición de 3.5 ml de isopropanol y agitación suave durante 15 segundos aproximadamente. Finalmente el DNA se recupera con una varilla de vidrio y se coloca en un tubo Eppendorf de 1.5 ml, evitando el traspaso excesivo de isopropanol, con 250 μ l de buffer de elución, un buffer de Tris-Cl con pH cercano a 8 que permite la solubilización del DNA (obtenido de Qiagen); posteriormente, se deja a temperatura ambiente durante 15-20 minutos, se retira la varilla con precaución y se deja resuspender en el buffer durante 1 hora a 50 °C con agitación moderada o a temperatura ambiente toda la noche, finalmente se mide la concentración.

Para la obtención de DNA genómico se utilizaron aproximadamente 5×10^6 - 1×10^7 células. Se puede realizar directamente con muestras frescas o en caso contrario congelar el pellet a -80 °C y descongelar en hielo antes de empezar la extracción.

4.6.- Cuantificación de ácidos nucleicos.

La cuantificación de ácidos nucleicos se realiza para conocer tanto la cantidad como la calidad del DNA obtenido en el proceso de extracción, se utiliza tanto para las muestras de DNA genómico como para las muestras de DNA plasmídico.

Para ello, utilizamos el aparato "*nanodrop 1000*", obtenido de Thermo, y el software incluido con él, que permite, mediante espectrofotometría y el uso de tan sólo 1 μ l de muestra, la cuantificación y la medida de calidad del DNA.

Brevemente, se coloca 1 μ l de agua destilada (en el pedestal inferior) para limpiar el instrumento se limpian los pedestales con papel de filtro y se coloca 1 μ l de agua o del buffer empleado para la elución del DNA, según corresponda, para realizar el blanco. Posteriormente se limpian de nuevo los pedestales y se coloca 1 μ l de la muestra para su medida; los pedestales deben limpiarse entre muestra y muestra. Al terminar las muestras se realiza una última medición con agua destilada de nuevo, para prolongar la vida útil del aparato.

La concentración de ácido nucleicos se realiza mediante la medición de la absorbancia de luz ultravioleta a 260 nm de longitud de onda, dándonos una densidad óptica (OD) que se traduce en mg/ml. Además de la concentración de DNA contenida, el Nanodrop nos proporciona una estimación de la pureza del DNA comparando la absorbancia a 260 nm frente a la absorbancia a 280 nm o la absorbancia a 230 nm, mediante los valores numéricos de los ratios 260/280 y 260/230; cuando se realiza una medición de DNA, el primero debe situarse cercano al 1.8, los valores inferiores indican contaminaciones con alguna solución, y los valores superiores indican posible contaminación con RNA.

A pesar de todo, y en el caso de DNA que va a ser utilizado para la transfección, es conveniente comprobar también el grado de superenrollamiento de los plásmidos, esto se utiliza como medida adicional de la calidad del DNA, ya que los plásmidos en estado superenrollado entran en las células en una proporción mucho más alta que los que han perdido

superenrollamiento, brevemente el proceso consiste en:

- Se elige la enzima de restricción a utilizar; dicha enzima debe cortar idealmente una vez en el vector, aunque puede cortar más de una vez; típicamente, para vectores basados en pCEP4 se utiliza la enzima *Bgl* II, que corta 2 veces en el esqueleto de dicho vector, liberando una banda de 9.5 kb y otra de 600 pb en el caso de vectores vacíos.

- Se coloca aproximadamente 1 µg del plásmido a comprobar (típicamente 2-3 µl dependiendo de la cantidad obtenida en la midiprep) con 1ml del buffer 10X correspondiente de la enzima de restricción (buffer 3 de New England Biolabs en el caso de *Bgl* II) y 0.1 µl de BSA (*Bovine Serum Albumin*) si la enzima lo requiere, y se lleva, generalmente, a 9.7 µl con agua destilada.

- Se añade la cantidad apropiada de enzima de restricción, generalmente 0.3 µl, y se incuba a 37 °C durante 1 hora en un termobloque, sin necesidad de agitación. El tiempo de incubación puede variar de unas enzimas a otras.

- Se comprueba la movilidad electroforética del DNA fragmentado frente a DNA sin fragmentar; el DNA sin digerir debe presentar una movilidad electroforética mucho mayor que el DNA digerido, indicando que se encuentra en estado superenrollado.

Una vez comprobada la calidad del DNA obtenido mediante “midiprep”, éste se puede utilizar para transfectar las células.

4.7.- Transfección de vectores.

Una vez se ha comprobado la calidad de los plásmidos, éstos se pueden emplear para realizar la transfección.

Todos los pasos para la preparación de la mezcla de transfección y la transfección de las células se realizaron en cabina de bioseguridad tipo II. Típicamente los experimentos se llevan a cabo en triplicados utilizando placas de 6 pocillos, con un volumen final de 100µl de mezcla de transfección por pocillo, brevemente el proceso consiste en:

Preparación de la mezcla de transfección.

- En un tubo de 1.5 ml estéril se añade 1 µg del DNA por pocillo a transfectar, hasta un máximo de 10 µg por tubo, y teniendo en cuenta añadir 1 µg de exceso para evitar problemas derivados de errores de pipeteo.
- En un tubo distinto de 1.5 ml estéril, se añaden 97 µl de Optimem y 3 µl del reactivo de transfección correspondiente (FuGENE 6 o X-tremeGENE 9 según la línea celular) por cada µg de DNA a transfectar, hasta un máximo de 10 µg por tubo (y teniendo en cuenta el exceso). Primero se añade el Optimem y luego se añade el reactivo de transfección en el centro del Optimem, con cuidado de no tocar las paredes del tubo; inmediatamente después se agita con el dedo 2-3 veces el tubo para mezclar el contenido, y se incuba a temperatura ambiente durante 5 minutos.
- Se transfiere la mezcla de reactivo de transfección y Optimem al tubo que contiene el DNA, evitando tocar

las paredes del tubo, se agita con el dedo 2-3 veces el tubo para mezclar el contenido, y se incuba esta mezcla durante 15-20 minutos a temperatura ambiente.

Transfección.

- En el caso de la transfección de DNA no es necesario quitar el medio previamente; se transfieren gota a gota 100 μ l de la mezcla de transfección por pocillo, intentando distribuir de manera homogénea.
- Al devolver la placa al incubador se agita con delicadeza 2-3 veces adelante-atrás y 2-3 veces a izquierda-derecha, para asegurar que la distribución de la mezcla es homogénea.
- Se incuban las células con la mezcla de transfección durante al menos 6 horas, pudiendo transcurrir hasta 24 horas, tras lo cual se reemplaza el medio por medio fresco para eliminar los restos de la mezcla de transfección.

4.8.- Preparación y transfección de RNAs de interferencia.

Todos los pasos necesarios para la preparación y la transfección de los siRNAs se realizaron en cabina de bioseguridad tipo II.

Para la resuspensión de los siRNAs se empleó el buffer específico de siRNAs de Dharmacon (Thermo), número de referencia NC9920186, diluido 1/5 con agua destilada libre de DNAsas y RNAsas, obtenida de Gibco, número de referencia 10977-023.

Las soluciones *stock* de siRNAs se realizaron a 25 μ M y se conservaron a -20 °C para su uso (ver más

abajo) o a $-80\text{ }^{\circ}\text{C}$ para su almacenamiento a largo plazo.

Para obtener una disminución de al menos un 90% en los niveles de expresión de la proteína diana, la transfección de los RNAs de interferencia (siRNAs) se debe realizar utilizando una concentración final de 20-100 nM del siRNA.

Para nuestros experimentos, y tras un proceso de optimización seleccionamos 40 nM como concentración final de siRNA para disminuir eficientemente la expresión de la proteína diana.

En el caso de la transfección de siRNAs en placas de 6 pocillos se utiliza un volumen final de 1 ml de mezcla de transfección por pocillo; la transfección se realiza en duplicado (dos pocillos por cada siRNA).

Preparación de la mezcla de transfección.

Se realiza una solución de uso de siRNA a una concentración de $5\mu\text{M}$ a partir de la solución *stock* $25\mu\text{M}$; se diluye una cantidad de ésta 1/5 en buffer de siRNA 1X. Dicha solución se prepara en fresco cada vez que se vaya a realizar una transfección. Cada alicuota de la solución *stock* no debe ser descongelada más de 3 veces, para evitar la pérdida de calidad o degradación del siRNA debido a los procesos de congelación/descongelación.

A continuación se describe el proceso, indicando las cantidades utilizadas por pocillo, el OptiMEM puede ser sustituido por medio de cultivo sin P/S (Penicilina/Estreptomicina) ni suero (FBS).

- En un tubo de 1.5 ml se colocan 8 μl de siRNA de la solución de uso ($5\mu\text{M}$) por pocillo a transfectar, se

añaden 92 μ l de Optimem, y se pipetea 2-3 veces el contenido evitando la formación de burbujas.

- Paralelamente, en un tubo aparte, se añaden 97 μ l de Optimem, y se colocan 3 μ l de reactivo de transfección DharmaFECT 4 en el centro del Optimem, con cuidado de no tocar las paredes del tubo. Inmediatamente después se repipetea 2-3 veces el contenido, evitando la formación de burbujas, y se incuba a temperatura ambiente durante 5 minutos.

- Se transfiere la mezcla de DharmaFECT4 al tubo que contiene el siRNA, evitando tocar las paredes del tubo, se repipetea 2-3 veces y se incuba esta mezcla durante 15-20 minutos a temperatura ambiente.

- Se retira el medio del pocillo que contiene las células mediante aspiración, con delicadeza y evitando a toda costa que el pocillo se quede completamente seco; inmediatamente después, se añaden 800 μ l de medio de cultivo al tubo que contiene la mezcla de transfección, se repipetea 2-3 veces y se transfiere 1 ml resultante con delicadeza directamente al pocillo. El medio de cultivo debe ser el mismo utilizado para el crecimiento de la línea celular a transfectar, pero no debe contener P/S, ya que estos antibióticos afectan seriamente a la viabilidad celular durante la transfección.

- Se incuban las células con la mezcla de transfección durante al menos 6 horas, pudiendo transcurrir hasta 24 horas si la viabilidad celular lo permite, tras lo cual se reemplaza el medio por medio fresco, que ya sí puede contener P/S.

4.9.- Cotransfección de vectores y RNAs de interferencia.

Todos los pasos necesarios se realizaron en cabina de bioseguridad tipo II. La cotransfección de siRNAs y DNAs se realiza de una manera casi idéntica a la transfección de los siRNAs en solitario, con la excepción del reactivo de transfección utilizado y la adición del DNA en el proceso. La concentración final de siRNA para disminuir eficientemente la expresión de la proteína diana se mantiene a 40 nM.

En el caso de la co-transfección de siRNAs y DNAs en placas de 6 pocillos se utiliza un volumen final de 1 ml de mezcla de transfección por pocillo, y la co-transfección se realiza por duplicado.

Preparación de la mezcla de co-transfección.

- Se realiza la solución de uso de siRNA a una concentración de 5 μ M a partir de la solución *stock* 25 μ M, preparada en fresco cada vez que se vaya a realizar una transfección o co-transfección. Cada alicuota de la solución *stock* no debe ser descongelada más de 3 veces, para evitar la pérdida de calidad o degradación del siRNA debido a los procesos de congelación/descongelación.

- Se prepara una solución de uso de DNA para la transfección con el fin de evitar errores debidos al pipeteo de pequeñas cantidades; brevemente, se diluye el DNA en una cantidad adecuada de Optimem, generalmente entre 1/5-1/10, para obtener una concentración de DNA en torno a 100 ng/ μ l. Al igual que la solución de uso de siRNA, esta solución debe

prepararse en fresco cada vez que se vaya a realizar la cotransfección.

- En un tubo de 1.5 ml se colocan 8 μ l de siRNA de la solución de uso (5 μ M) por pocillo a transfectar, se añaden los ml de Optimem necesarios para llegar a 100 μ l de volumen final (habitualmente en torno a 82 μ l), y por último se añade 1 μ g del DNA a transfectar (aproximadamente 10 μ l de la solución de DNA diluido), la mezcla se repipetea 2-3 veces evitando la formación de burbujas.

- Paralelamente, en un tubo aparte, se añaden 97 μ l de Optimem, y se colocan 3 μ l de reactivo de transfección DharmaFECT DUO en el centro del Optimem, con cuidado de no tocar las paredes del tubo, inmediatamente después se repipetea 2-3 veces el contenido, evitando la formación de burbujas a toda costa, y se incuba a temperatura ambiente durante 5 minutos.

- Se transfiere la mezcla de DharmaFECT DUO al tubo que contiene la mezcla de siRNA-DNA, evitando tocar las paredes del tubo, se repipetea 2-3 veces y se incuba esta mezcla durante 15-20 minutos a temperatura ambiente.

Co-transfección.

- Se retira el medio del pocillo que contiene las células mediante aspiración, con delicadeza y evitando a toda costa que el pocillo se quede completamente seco; inmediatamente después, se añaden 800 μ l de medio de cultivo al tubo que contiene la mezcla de transfección, se repipetea 2-3 veces y se transfiere el ml resultante con delicadeza directamente al pocillo.

El medio de cultivo debe ser el mismo utilizado para el crecimiento de la línea celular a transfectar, pero no debe contener P/S, ya que estos antibióticos afectan seriamente a la viabilidad celular durante la transfección.

- Se incuban las células con la mezcla de transfección durante al menos 6 horas, pudiendo transcurrir hasta 24 horas si la viabilidad celular lo permite, tras lo cual se reemplaza el medio por medio fresco, que ya sí puede contener P/S.

4.10.- Ensayos en líneas celulares.

4.10.1.- Ensayo de resistencia a mitomicina C.

Se siembra en duplicados o triplicados una cantidad adecuada de células: 1×10^4 células por pocillo para líneas de fibroblastos humanas en placas de 6 pocillos o 5×10^3 células por pocillo en placas de 12 pocillos; en el caso de líneas celulares de hámster, se siembran 5×10^3 células por pocillo en placas de 6 pocillos.

El día posterior a la siembra se cambia el medio por medio con diferentes concentraciones de MMC (0 nM, 50 nM o 100 nM). Dos días después se vuelve a cambiar el medio y a las 72 horas de empezar el tratamiento con MMC se fijan las células (figura 24).

Fijación.

Para fijar las células para su posterior tinción, se aspira el medio de cultivo de la placa y se añade una cantidad adecuada de solución fijadora de ácido acético y metanol (composición en materiales), típicamente 500 μl por pocillo en placas de 12 pocillos o 1 ml por pocillo en placas de 6 pocillos, y se incuba durante 10 minutos a temperatura ambiente con agitación.

Posteriormente se aspira la solución de fijación, se realizan dos lavados con agua destilada y se procede directamente a la tinción de las colonias.

Revelado por tinción con solución de cristal violeta.

Se desecha el agua destilada de los pocillos y se añade una cantidad apropiada de solución de cristal violeta, aproximadamente 500 μl por pocillo en placas de 12 pocillos o 1ml por pocillo en placas de 6 pocillos, y se incuba a temperatura ambiente durante 10 minutos, permitiendo la tinción de las colonias resultantes del experimento (revelado).

Cuantificación con solución de disolución.

Tras la tinción de las colonias, se retira la solución de cristal violeta y se realizan dos lavados de las placas con agua destilada y se dejan secar las placas hasta el día siguiente; posteriormente, se añade una cantidad apropiada de solución de disolución, concretamente 250 μl por pocillo en placas de 12 pocillos o 500 μl por pocillo en placas de 6 pocillos, y se incuba a temperatura ambiente con agitación durante 2 horas; posteriormente, se

transfieren 100 μ l (placas de 12 pocillos) o 200 μ l (placas de 6 pocillos) de cada pocillo a un pocillo de una placa de 96 pocillos y se mide la densidad óptica mediante un lector de placas, con el ajuste OD₅₉₅, en nuestro caso se empleó el GloMax-Multi de Promega. Se establece como 100% el valor obtenido en las células sin tratar y se calcula el porcentaje de resistencia a MMC en comparación con este valor.

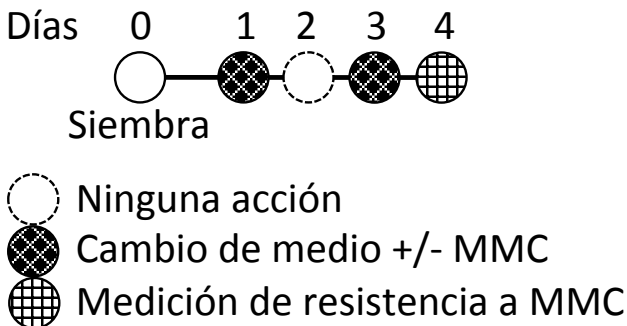


Figura 24. Esquema del ensayo de resistencia a MMC a 72 horas en células en cultivo. Los círculos indican las acciones a realizar cada día.

4.10.2.- Ensayo de retrotransposición.

Todos los ensayos de retrotransposición se realizaron de forma transiente (Wei et al, 2000) analizando la retrotransposición en células mediante la expresión de un gen reportero (figura 25). Brevemente, se clona aguas abajo del retroelemento un casete que contiene un gen reportero, que se encuentra antisentido con respecto a la dirección transcripcional del elemento LINE-1; este gen contiene su propio promotor y señal de poliadenilación, pero está interrumpido por un intrón en la misma dirección transcripcional que el

retroelemento, de manera que la expresión sólo puede producirse tras:

- La eliminación del intrón que interrumpe el gen por *splicing* del RNA transcrito aguas abajo del retroelemento y,
- Su posterior integración en el genoma mediante retrotransposición.

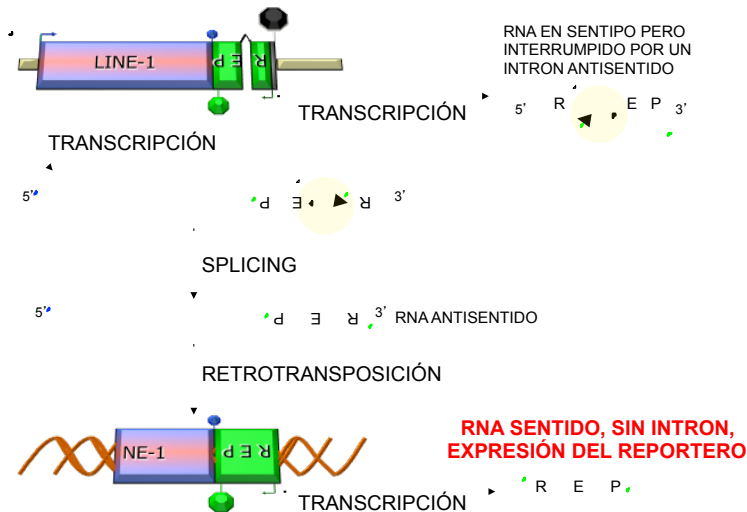


Figura 25. Esquema del funcionamiento del ensayo de retrotransposición en células en cultivo. El gen reportero sólo puede expresarse desde una localización genómica tras un auténtico evento de retrotransposición.

Para llevar a cabo los ensayos de retrotransposición, se siembra un número adecuado de células por pocillo, en función de la línea celular estudiada; típicamente 1×10^4 células por pocillo en placas de 6 pocillos; posteriormente, 16-18 horas tras

la siembra, se realiza la transfección con el reactivo de transfección adecuado; al día siguiente se cambia el medio para retirar los restos de la transfección.

De 3 a 5 días después de la transfección se realiza la selección con el antibiótico correspondiente, manteniéndose 12 días en el caso de la neomicina (G418) o durante 8 días en el caso de la blasticidina. Transcurrido ese tiempo una célula en la que se ha producido un evento de retrotransposición ha formado colonias que se puede fijar y teñir (figura 26).

Posteriormente se cuentan las colonias resultantes y se normaliza el resultado en función de la eficiencia de transfección obtenida para el plásmido correspondiente, descrito a continuación.

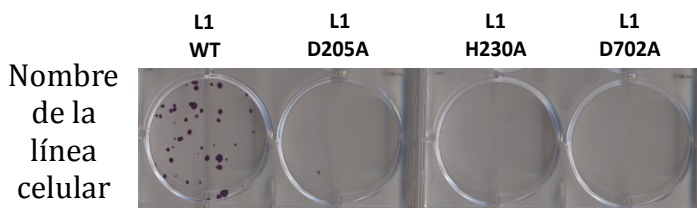


Figura 26. Visualización o revelado de un experimento típico de retrotransposición con casete de resistencia a antibiótico. En la fila superior se indica el elemento utilizado, a la izquierda la línea celular en la que se realiza el experimento; de izquierda a derecha, se representan los elementos salvaje (WT), mutantes en el dominio endonucleasa (D205A y H230A), y el mutante en el dominio reversotranscriptasa (D702A). Cada colonia representa un nuevo evento de retrotransposición, en circunstancias normales la movilidad se produce al utilizar el elemento salvaje, mientras que en los mutantes de la endonucleasa se detecta poca o ninguna actividad, y ninguna actividad en los mutantes del dominio reversotranscriptasa.

Eficiencia de transfección.

En los casos en los que se requiere comparar la actividad de los diferentes mutantes de LINE-1 entre sí, es necesario normalizar el número de colonias obtenidas en función de la eficiencia de transfección del plásmido utilizado; para ello, y a pesar de que normalmente la eficiencia con la que los diferentes plásmidos utilizados en el mismo experimento son incluidos en las células es muy similar, es necesario realizar un experimento control que mida la eficiencia con la que cada plásmido es incluido en las células transfectadas.

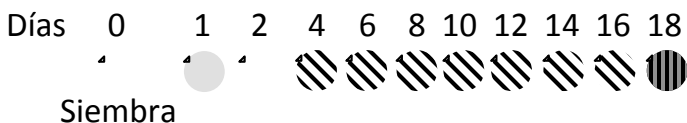
Para ello, se realiza una cotransfección con cantidades equivalentes de constructo de expresión de LINE-1 y un vector de expresión de un gen reportero, en nuestro caso dicho gen es la proteína verde fluorescente (GFP) contenida en el vector de expresión eucariota pCEP4. Cada una de las células sometidas a transfección acepta ambos plásmidos o ninguno, de manera que se puede medir la proporción de células que ha sido realmente transfectada en el ensayo de retrotransposición por extrapolación con este valor; alternativamente, como en nuestros experimentos, se puede utilizar la diferencia en la eficiencia de transfección para establecer una proporción más exacta de las diferencias obtenidas en el recuento de colonias resultantes del ensayo de retrotransposición.

Para realizar el ensayo de eficiencia de transfección se siembra una cantidad adecuada de células, típicamente 1×10^5 células por pocillo en placas de 6 pocillos; 16-18 horas después se realiza la cotransfección con los plásmidos de expresión de GFP y de LINE-1, al día siguiente se cambia el medio para

eliminar los restos de la transfección y, 72 horas tras la transfección, se realiza un análisis de expresión de GFP mediante citometría de flujo para estimar la proporción de células positivas. Esto nos dará una estimación del porcentaje de transfección. Las diferencias obtenidas son después extrapoladas al recuento de colonias del ensayo de retrotransposición para establecer un porcentaje de actividad fidedigno del mutante estudiado con respecto al elemento salvaje.

4.10.2.1.- Selección por neomicina.

En los ensayos en los que los vectores de expresión de L1 transfectados llevan el casete del gen reportero de resistencia a neomicina (casete mneol, Freeman et al., 1994), a los 3 días de realizar la transfección (día 4 del experimento), se empieza la selección con la cantidad adecuada del antibiótico G418 (neomicina); dicha selección se realiza cambiando el medio por medio fresco con G418 cada dos días, manteniendo la selección durante 14 días (figura 27). Pasado ese tiempo se procede a revelar el ensayo mediante las técnicas de fijación y tinción descritas previamente.



- Cambio de medio
- Transfección del plásmido de expresión de L1
- ▨ Cambio de medio con G418
- Fijación/Tinción de colonias resultantes

Figura 27. Esquema del ensayo de retrotransposición con gen reportero de resistencia a neomicina. Los círculos indican las acciones a realizar en el día correspondiente.

4.10.2.2.- Selección por blasticidina.

En los ensayos en los que los vectores de expresión de L1 transfectados llevan el casete del gen reportero de resistencia a blasticidina (mblastI, Goodier et al., 2007 y Morrish et al., 2002), se cambia el medio por medio fresco sin antibiótico el día después de realizar la transfección y dos días después, esperando un total de 5 días desde la transfección antes de empezar la selección con una cantidad adecuada de blasticidina en función de la línea celular estudiada; para la selección con este antibiótico no es necesario realizar cambios de medio durante el ensayo, si bien se puede cambiar por medio fresco en el caso de que las células sean sensibles a la muerte celular de otras células. La selección por blasticidina se mantiene durante 8 días (figura 28), tras los cuales se procede a la fijación y tinción de las colonias resultantes.

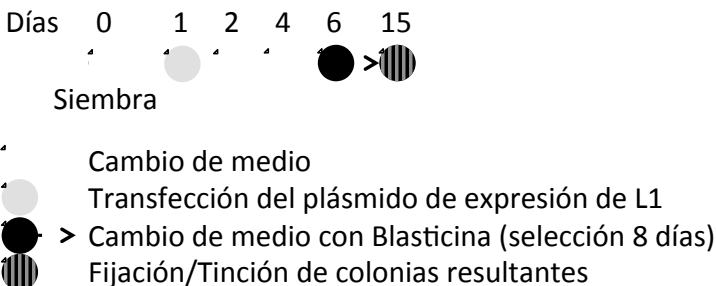


Figura 28. Esquema del ensayo de retrotransposición con gen reportero de resistencia a blasticidina. Los círculos indican las acciones a realizar en el día correspondiente.

4.10.2.3.- Con RNAs de interferencia.

En estos ensayos, primero se realizó una transfección del siRNA correspondiente en solitario, al día siguiente se cambió el medio por medio fresco y 48 horas después de la transfección del siRNA se resemebró un número apropiado de células; un día después, se realizó la cotransfección del vector de expresión de L1 con el siRNA correspondiente, de nuevo al día siguiente se cambió el medio por medio fresco para retirar los restos de la transfección y 72 horas después de la co-transfección se empezó la selección por blasticidina. En paralelo se realizó el mismo experimento pero transfectando el plásmido pCDNA6, vector de expresión del gen de resistencia a blasticidina sin intrón, para poder normalizar el número de colonias resultantes del ensayo de retrotransposición.

Como control de la actividad del siRNA se realizó en paralelo un experimento control de sensibilidad a mitomicina (figura 29); brevemente, el día después de la siembra de las células, se transfectó

el siRNA correspondiente, al día siguiente por la mañana se cambia el medio por medio fresco para retirar los restos de la transfección y unas 6 horas después se procede a resembrar las células, al día siguiente se cambia el medio por medio con mitomicina C, 48 horas después se cambia el medio de nuevo por medio con mitomicina C y al día siguiente se procede a revelar el ensayo, de manera muy similar al ensayo de resistencia a mitomicina C descrito previamente (Figura 24).

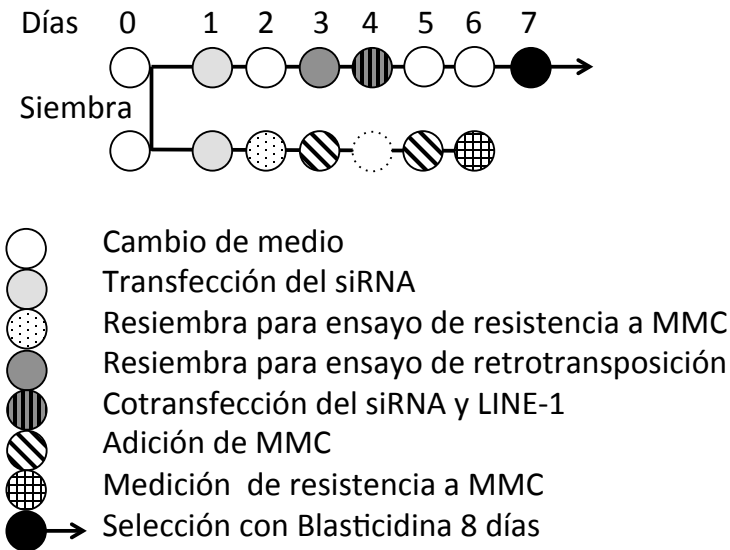


Figura 29. Esquema del ensayo de RNAs de interferencia y retrotransposición con gen reportero de resistencia a blasticidina. Los círculos indican las acciones a realizar en el día correspondiente, el ensayo de retrotransposición (fila superior) se realiza en paralelo con el ensayo control de resistencia a MMC (fila inferior).

4.11.- Recuperacion de inserciones de LINE-1.

Ensayo De Retrotransposición Con Selección por Neomicina.

Todos los ensayos de retrotransposición se realizaron siguiendo el esquema de ensayo descrito previamente (figura 27), utilizando el elemento LINE-1 salvaje o mutante en su dominio endonucleasa.

Para poder recuperar las inserciones de los elementos LINE-1, hemos utilizado un casete reportero modificado (Gilbert et al., Cell 2002). Brevemente, el casete consiste en un casete mneol modificado; este casete contiene un origen de replicación bacteriano (ColE1) corriente abajo del casete mneol; por otro lado, el gen NEO contiene un promotor procariota (EM7) y una secuencia Shine-Dalgarno para mediar su traducción en *E. coli*; con esta configuración, tras un evento de movilidad, el casete NEO puede dar resistencia a neomicina y kanamicina según el tipo celular.

Para llevar a cabo los ensayos de retrotransposición se siembra un número apropiado de células en función de la línea celular y el elemento utilizados:

- Células de hámster: en placas de cultivo celular de 100 mm se siembran 1×10^4 para elementos wild type o 5×10^4 para elementos que llevan la mutación D205A.
- Fibroblastos humanos inmortalizados: en placas de cultivo celular de 100mm se siembran 8×10^4 para elementos wild type o 2×10^5 para elementos que llevan la mutación D205A. Para mejorar la viabilidad

de las células tras la transfección se realiza un revestimiento de la placa, previo a la siembra, con una solución de gelatina al 0.1% p/v (1 g en 1 l de agua destilada). La gelatina procede de piel porcina, obtenida de Sigma-Aldrich, número de referencia G2500-500g. Brevemente, para realizar el revestimiento, se prepara una solución *stock* de gelatina al 1% y se autoclava, posteriormente se realiza una dilución 1/10 para obtener la solución de uso al 0.1% y se filtra con un filtro o una unidad de filtración con un tamaño de poro de 0,22 μm . Por último, se añaden 4 ml de gelatina al 0.1% p/v a la placa y se incuba a 37 °C en el incubador durante 2 horas, tras lo cual se aspira la gelatina y se siembran las células.

Cuando finaliza la selección, en lugar de fijar y teñir las colonias resultantes, se procede al aislamiento y establecimiento de líneas celulares clonales (figura 30), Alternativamente, se puede realizar el proceso con la población celular al completo.

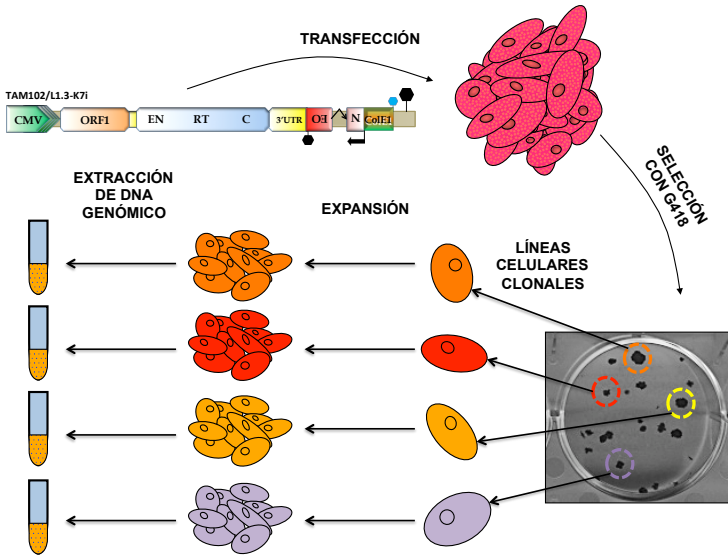


Figura 30. Obtención de líneas celulares clonales con inserciones de L1. Cada una de las líneas celulares aisladas resistentes a G418 alberga al menos un evento de retrotransposición único distinto de los presentes en las otras líneas celulares evitando elegir colonias que se encuentren muy próximas entre sí.

Aislamiento y establecimiento de Líneas clonales.

Tras finalizar la selección, se retira el medio, se lavan las células con PBS, se aspira el PBS y posteriormente, se realiza el aislamiento individual de 3-6 colonias resultantes de la selección.

Para la obtención de líneas clonales, mediante la ayuda de unas pinzas, se cubre la colonia con un pequeño disco de papel de filtro de Whatman, impregnado previamente en tripsina, evitando elegir una colonia que se encuentre próxima a otra.

Se deja actuar durante 5 minutos a 37 °C, y se pasa el papel a un pocillo de una placa de 6 pocillos

que contiene 2 ml de medio con neomicina, se agita el papel con ayuda de las pinzas y se deja en el pocillo hasta el día siguiente para asegurar la liberación del máximo número posible de células. En el caso de aislar varias colonias, es necesario introducir brevemente las pinzas en etanol y luego secar entre colonias para evitar contaminaciones cruzadas.

Expansión y congelación.

Las colonias aisladas se cultivan hasta alcanzar aproximadamente un 100% de confluencia, tras lo cual se tripsinizan y se dividen en 2 pocillos, cuando estos pocillos alcancen el 100% uno se tripsiniza para transferir las células a un frasco T75 y las células del otro pocillo se congelan en primer lugar a -80 °C y posteriormente en nitrógeno líquido, para evitar perder la línea.

Cuando las células en el frasco de T75 alcanzan el 100% de confluencia, se procede a la extracción de su DNA genómico (descrita anteriormente), y al rescate de las inserciones (figura 31).

Alternativamente, las células obtenidas se pueden lavar con PBS y almacenar a -80 °C para realizar la extracción de DNA posteriormente.

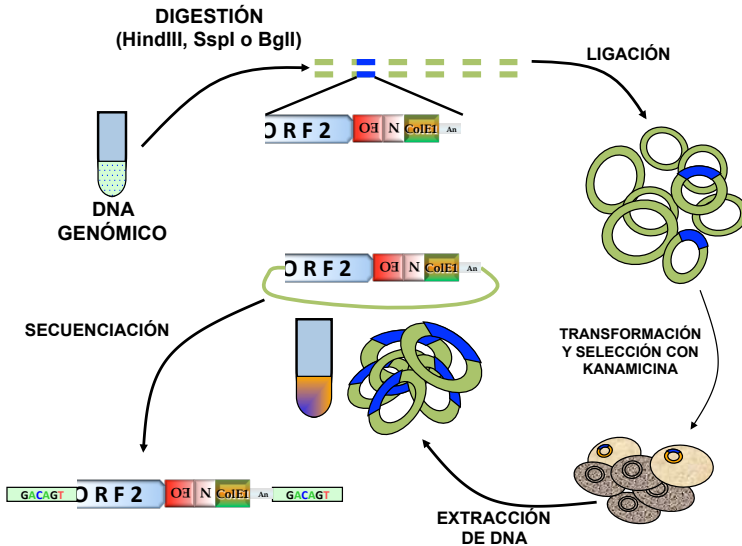


Figura 31. Rescate de inserciones de L1. El DNA genómico fragmentado y religado se transforma en células competentes, de manera que sólo los fragmentos que contienen al menos el cassette k7i pueden crecer. Tras el aislamiento de los clones obtenidos se empieza la secuenciación para caracterizar la inserción de L1.

Restricción y ligación intramolecular de fragmentos de DNA genómico.

Una vez obtenido el DNA genómico de las colonias de células aisladas se procede a digerirlo con una enzima de restricción que corta múltiples veces en el genoma humano, pero que no corta en nuestro casete reportero. Brevemente, se digieren 8 µg de DNA genómico con 25-100 unidades de enzima de restricción (un claro exceso molar), en un volumen total de 100 µl, que contiene 10 µl del buffer 10X correspondiente a la enzima de restricción utilizada y 1 µl de BSA 100x, y se incuba a 37 °C durante al menos 12 horas. Tras 12 horas, se añade 1 µl más de

enzima de enzima de restricción y se incuba a 37 °C durante 2 horas más.

Posteriormente, se inactiva la enzima de restricción por calor a 65 °C, o se elimina mediante purificación por columna en el caso de que ésta no sea sensible a calor.

Tras la inactivación de la enzima de restricción, se procede a la ligación de los fragmentos resultantes. Para favorecer la ligación intramolecular, se realiza en condiciones diluidas, llevando la reacción a un volumen total de 500 µl; se aumenta el volumen de agua 351 µl de destilada, 40 µl de buffer de ligasa 10X y aproximadamente 3200 unidades de T4 ligasa, en nuestro caso 8 µl de la enzima T4 ligasa de New England Biolabs, número de referencia M0202L, de 400 U/µl; las reacciones de ligación se incuban durante toda la noche a 16 °C. A la mañana siguiente, se añaden aproximadamente 600 unidades más (1.5 µl) de T4 ligasa y se incuba a temperatura ambiente durante 4 horas más.

Transformación de bacterias ultracompetentes.

Para obtener los plásmidos que contienen un evento de inserción de LINE-1 flanqueado por fragmentos de DNA genómico, primero es necesario concentrar el DNA resultante de la reacción de ligación.

Brevemente, se transfiere todo el volumen de la reacción de ligación a una unidad de filtración de AmiconYM-100 (obtenidas de Millipore) y se centrifuga a 5000 x g durante 5 minutos a temperatura ambiente. Posteriormente se desecha el contenido del tubo de colección y se añaden 500 µl de

agua destilada a la columna de la unidad, se centrifuga de nuevo a 5000 x g durante 5 minutos y se repite el proceso.

Tras los dos lavados con agua destilada para eliminar sales que puedan interferir con la transformación, se inserta la columna de la unidad de filtración en posición invertida en un nuevo tubo y se centrifuga brevemente a máxima velocidad, obteniéndose generalmente un volumen de unos 30-50 μ l.

Posteriormente se realiza la transformación en bacterias ultracompetentes XL10-Gold, en las condiciones descritas anteriormente. Brevemente, se utiliza parte de la ligación concentrada (10 μ l), a los que se añaden 500 μ l de bacterias ultracompetentes, y se incuban en hielo durante 2 horas en tubos de 15 ml de polipropileno.

Tras la incubación, se realiza el choque térmico a 42 °C durante 38 segundos con ligera agitación y se transfieren los tubos a hielo durante 2 minutos. Posteriormente se añade 1 ml de LB sin antibiótico a cada tubo y se recupera a las bacterias durante 12 horas a temperatura ambiente con agitación, dejando el tubo sin cerrar completamente para permitir la aireación de los cultivos.

Tras la larga recuperación, se centrifugan las bacterias a 500 x g durante 5 minutos, y se retira cuidadosamente el sobrenadante, posteriormente se resuspenden las bacterias en 500 μ l de LB sin antibiótico, de los cuales:

- Se siembran 400 μ l en placas con kanamicina de 140mm de diámetro,

- Se transfieren los 100 μ l restantes a un tubo con 2 ml de LB líquido con kanamicina, en el que se crecen durante 24 horas más a 37 °C con agitación. Pasado ese tiempo, se centrifugan las bacterias a 500 x g durante 5 minutos, se retira cuidadosamente el sobrenadante y se resuspenden las bacterias en 100 μ l de LB sin antibiótico, para posteriormente sembrarlas en una placa con kanamicina de 90 mm de diámetro.

Posteriormente, se aíslan colonias resistentes a kanamicina y se extrae el DNA plasmídico. Finalmente, se hace un primer chequeo con la enzima de restricción *Eco* RI para comprobar la presencia del plásmido, tras el cual se puede empezar la caracterización de la inserción.

En el caso de no obtener transformantes tras la primera transformación, se repite el proceso con el resto de la ligación.

4.12.- Caracterización de las inserciones recuperadas.

En la primera ronda de secuenciación se utilizan dos oligonucleótidos diferentes en direcciones opuestas, uno que parte de la secuencia de origen de replicación bacteriano hacia la cola de poliA de la inserción, llamado "Reco3" y otro que parte desde el gen de neomicina en dirección al elemento LINE-1, llamado "Neo210as" (descritos en la sección de materiales 3.10). Posteriormente se van utilizando oligonucleótidos situados en el L1 en dirección 5' para determinar el punto de truncación del elemento así como su localización genómica, y se diseñan otros en dichas zonas si es necesario (descritos en la sección de materiales 3.10).

Una vez obtenidas las secuencias se alinean con el cassette de resistencia a neomicina y con el retroelemento, hasta determinar su localización genómica.

Herramientas informáticas para análisis (Software).

Serial Cloner

Este programa de *software* libre permite el alineamiento de dos secuencias mediante blast, enfrentando la secuencia obtenida de la secuenciación con la del constructo empleado para obtener las características de la inserción. Una vez que se alcanza el DNA genómico se utiliza el software correspondiente para determinar la localización y modificaciones producidas en el sitio de inserción.

BLAT

También conocido como *Genome Browser*, desarrollado por la Universidad de California en Santa Cruz (UCSC), permite determinar la localización genómica de la secuencia introducida en el genoma humano, si bien contiene genomas de múltiples organismos (Kent, 2002; Kent et al., 2002).

CHO Genome Database

El *Genome Browser* no contiene, en el momento de analizar los resultados este estudio, el genoma de Hámster o de las células CHO, y de hecho actualmente no existe ninguna construcción completa ensamblada del genoma de estas células; no obstante, en el año 2012 se formó un grupo que creó la web www.chogenome.org, actualmente mantenida por la Universidad de Delaware, con datos de la secuenciación del genoma de las células CHO-K1 (Xu et al., 2011), por lo que pudimos enfrentar nuestras secuencias con las piezas de genoma secuenciado contenidas en esta web, pudiendo determinar las características de la inserción. A pesar de ello, aún no se pueden ubicar en una localización cromosómica concreta, al no estar disponible el ensamblaje completo.

RESULTADOS

5.- RESULTADOS.

Con el fin de determinar si la actividad de L1 guarda relación con la sintomatología propia de los pacientes de anemia de Fanconi, así como para discernir si existe alguna relación entre la movilidad del retroelemento L1 y la maquinaria de reparación y replicación celular, se realizaron múltiples experimentos, que se describirán en los siguientes apartados, diseñados para medir el nivel de actividad de L1 ante deficiencias de las proteínas implicadas en estos procesos celulares.

5.1.- Ensayos de retrotransposición en células de hámster deficientes en la ruta de Fanconi.

Para determinar si la retrotransposición de LINE-1 mediante la vía independiente de la endonucleasa (ENi) tiene lugar en células deficientes para la ruta de reparación de la anemia de Fanconi (FA), se realizaron ensayos de retrotransposición con una batería de vectores de expresión de L1 y posterior selección con blasticidina. Se emplearon células de hámster deficientes en los miembros FANCG, BRCA2, FANCA y FANCC de la ruta de reparación FA.

5.1.1.- Ensayos de retrotransposición en células deficientes en FANCG.

En primer lugar se analizó la línea deficiente en FANCG CHO KO40 (figura 32), así como la línea parental de ésta, 40B6, que no presenta ninguna deficiencia en mecanismos de reparación (Tebss et al., 2005). Como control negativo para la retrotransposición se utilizó el mutante D702A, que contiene una mutación puntual en el dominio reversotranscriptasa (RT) de LINE1, que anula completamente la actividad de ésta sin afectar a la ribonucleopartícula del elemento (Doucet et al., 2010).

La movilidad del mutante D702A fue nula tanto en la línea parental como en la línea mutante para FANCG. Asimismo, tanto en la línea parental 40B6 como en la línea mutante KO40, observamos la aparición de algunas colonias resistentes a blasticidina en las células transfectadas con elementos mutantes en el dominio endonucleasa. Los porcentajes de movilidad alcanzados fueron de hasta un 6% de movilidad en comparación con los elementos L1 salvajes (datos corregidos con respecto a la eficiencia de transfección), no apreciándose diferencias en el nivel de movilidad de los mutantes en el dominio endonucleasa (ENi) en células deficientes en FANCG (figura 32).

Por tanto, las células de hámster deficientes en FANCG se comportan como células salvajes o *wild type* y no se observan niveles de movilidad ENi significativos.

A)



B)

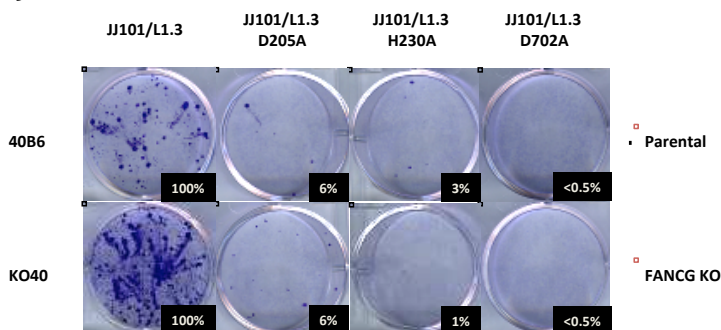


Figura 32. Ensayo de retrotransposición en células 40B6 y la línea derivada 40B6, deficiente en FANCG. A) Esquema del vector utilizado en los experimentos, señalando la posición relativa de las mutaciones en la secuencia del elemento. **B)** Tinción con cristal violeta de los resultados del experimento de retrotransposición. Se aprecia un nivel bajo de retrotransposición por la vía ENi en ambas líneas (1-6%), al comparar con el nivel de retrotransposición de los elementos salvajes en la línea correspondiente (datos corregidos según eficiencia de transfección).

5.1.2.- Ensayos de retrotransposición en células deficientes en FANCC.

Se ensayó la movilidad del elemento, también con la batería de vectores de expresión de L1 que permiten la selección por blasticidina, en la línea deficiente en FANCC (CLV5B), así como en la línea parental de ésta, V79B, que no presenta ninguna deficiencia en mecanismos de reparación de DNA (figura 33).

La movilidad del mutante de la RT de LINE1, así como la de los dos mutantes de la endonucleasa del elemento, fue nula o muy baja en la línea parental, mientras que la línea deficiente CLV5B presentó unas tasas de movilidad para los mutantes D205A y H230A del 25% y el 12%, respectivamente, al compararla con el elemento salvaje (corregidos con la eficiencia de transfección). El control negativo D702A, como se esperaba, no presentaba movilidad alguna (figura 33).

Por tanto, las células de hámster deficientes en FANCC se caracterizan por albergar niveles elevados de retrotransposición por la vía ENi.

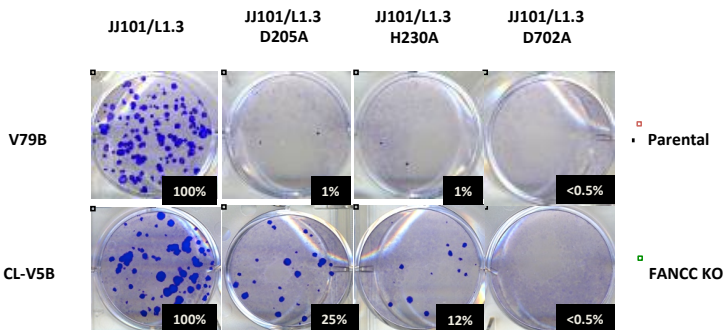


Figura 33. Ensayo de retrotransposición en células V79B y la línea derivada CLV5B, deficiente en FANCC. Tinción con cristal violeta de los resultados del ensayo de retrotransposición. La línea parental presenta un nivel bajo de retrotransposición por la vía ENi (1%), mientras que la línea CLV5B muestra niveles de actividad de los mutantes en el dominio endonucleasa D205A y H230A de 25% y 12%, respectivamente, en comparación con el nivel de retrotransposición de los elementos salvajes (datos corregidos según el nivel de eficiencia de transfección de los distintos plásmidos).

5.1.3.- Ensayos de retrotransposición en células deficientes en FANCA.

Paralelamente, se realizó el ensayo de retrotransposición, con la batería de vectores de expresión de L1 con selección por blasticidina, en una línea celular que presenta una mutación en el gen FANCA, concretamente en la línea VH4, así como en la línea parental de ésta, V79, que no presenta ninguna deficiencia en mecanismos de reparación (figura 34).

La movilidad del control negativo D702A y la de los mutantes en la endonucleasa del elemento fue nula o muy baja en la línea parental V79, mientras

que la línea VH4 los niveles de movilidad de los mutantes D205A y H230A alcanzaron el 62,5% y el 21,5%, respectivamente, en comparación con el elemento salvaje (corregido en función de la eficiencia de transfección), el control negativo D702A, como se esperaba, no presentó movilidad alguna (figura 34).

Por tanto, las células de hámster deficientes en FANCA se caracterizan por albergar niveles elevados de retrotransposición por la vía ENi.

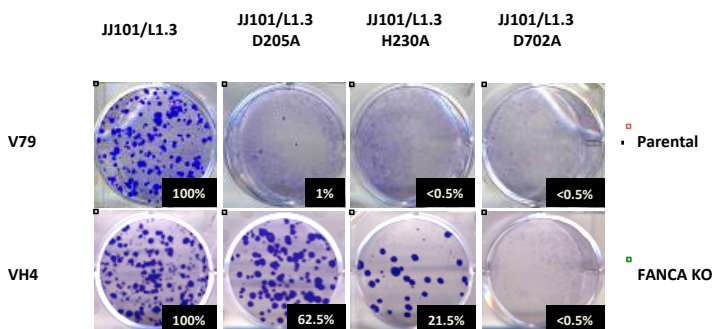


Figura 34. Ensayo de retrotransposición en células V79 y la línea derivada VH4, deficiente en FANCA. Los niveles de retrotransposición observados en la línea parental V79 se ajustan a lo esperado en células normales; en cambio, la línea VH4 deficiente en FANCA muestra niveles muy elevados de movilidad de los mutantes en el dominio endonucleasa D205A y H230A (62,5% y 21,5%, respectivamente), en comparación con el nivel de retrotransposición de los elementos salvajes (datos corregidos según eficiencia de transfección).

5.1.4.- Ensayos de retrotransposición en células deficientes en FANCD1/BRCA2.

También se realizó el ensayo de retrotransposición con la batería de vectores de expresión de L1 con selección por blasticidina en la misma línea parental, V79, y la línea derivada VC8, que presenta una mutación en el gen BRCA2 (figura 35).

La movilidad del control negativo D702A fue nula en ambas líneas, así como la de los mutantes en la endonucleasa del elemento L1 en la línea parental V79; de forma similar, los niveles de retrotransposición de los mutantes en el dominio endonucleasa, inferiores al 5%, no presentaron una movilidad significativa en comparación con el elemento salvaje en la línea VC8, deficiente en el gen BRCA2 (figura 35).

Por tanto, las células de hámster deficientes en FANCD1 se comportan como células salvajes o *wild type* y no se observan niveles de movilidad ENi significativos.

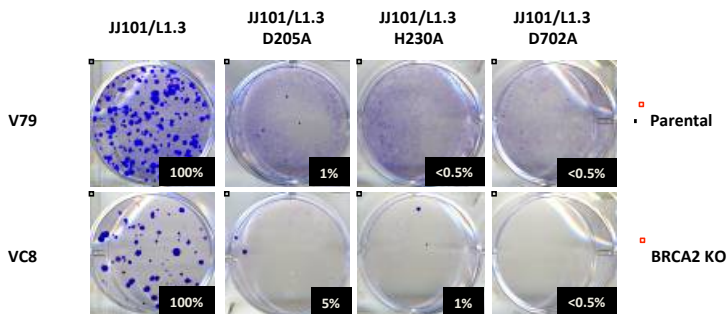


Figura 35. Ensayo de retrotransposición en células V79 y la línea derivada VC8, deficiente en BRCA2 (FANCD1). El ensayo de retrotransposición en la línea VC8, deficiente en BRCA2/FANCD1, no mostró un incremento significativo de la movilidad por la vía ENi, comparado con el nivel de retrotransposición de los elementos salvajes (datos corregidos según eficiencia de transfección).

5.2.- Ensayos de retrotransposición en células de hámster complementadas.

A continuación, para comprobar si el aumento de la movilidad por la vía independiente de la endonucleasa de los retroelementos ensayados se debe a la deficiencia del gen de Fanconi afectado en cada una de las líneas mutantes y no a otras causas, se realizaron ensayos de complementación para cada uno de los genes estudiados.

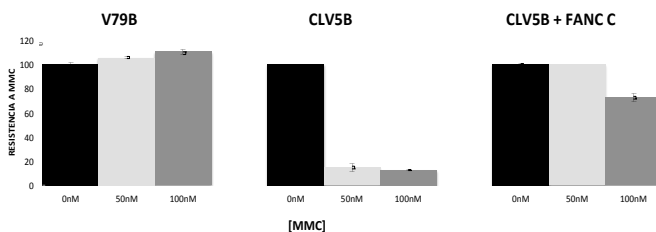
5.2.1.- Ensayos de retrotransposición en células CLV5B complementadas con el gen humano FANCC.

El primero de los genes a estudiar fue FANCC. Realizamos ensayos de complementación de la línea CLV5B con el gen humano de FANCC mediante infección por retrovirus, y sometimos a las células a un nuevo experimento de retrotransposición con la batería de vectores de expresión de LINE1 que incluían el *cassette* de selección por blasticidina (figura 36B).

La movilidad del mutante de la reversotranscriptasa de LINE1 que contiene la mutación puntual D702A fue nula en la línea complementada, y los niveles de los mutantes de endonucleasa en esta línea fueron nulos (H230A) o muy bajos (D205A, 4.8%) (figura 36B).

Para comprobar que la introducción del gen FANCC corregía también la sensibilidad a agentes intercalantes se sometieron las tres líneas celulares a un ensayo de resistencia a mitomicina C (MMC) (figura 36A). La línea parental, V79B, no presentaba sensibilidad alguna, incluso a altas dosis de MMC, mientras que la línea deficiente en FANCC CLV5B, presentó una sensibilidad extrema al agente, incluso a una dosis relativamente baja de MMC, 50nM. Por el contrario, la línea CLV5B complementada con el gen FANCC recuperó completamente la resistencia a MMC a 50nM. A una dosis de 100nM, la sensibilidad se encontró parcialmente corregida, alcanzando casi un 75% de resistencia con respecto a las células control sin tratar (figura 36A).

A)



B)

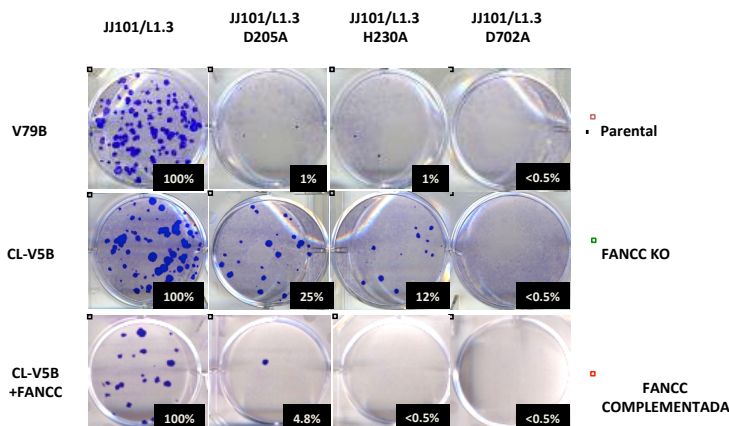


Figura 36. Complementación de la línea CLV5B con el gen humano FANCC. A) **Ensayo de resistencia a Mitomicina C.** La gráfica muestra el porcentaje de células vivas tras el tratamiento a las distintas dosis indicadas de mitomicina C (MMC) en comparación con las células sin tratar (barra en negro, 0nM). La línea complementada con FANCC muestra una corrección total de la sensibilidad a MMC a una dosis de 50nM y parcial (alrededor del 70%) a una dosis de 100nM. B) **Ensayo de retrotransposición.** La complementación de la línea CLV5B con el gen humano de FANCC es capaz de corregir casi por completo las diferencias en los niveles de movilidad de los mutantes en el dominio endonucleasa de L1, disminuyendo los niveles de retrotransposición del mutante D205A hasta el 4,8% y no detectándose en el mutante H230A (datos corregidos según eficiencia de transfección).

5.2.2.- Ensayos de retrotransposición en células VH4 complementadas con el gen humano FANCA.

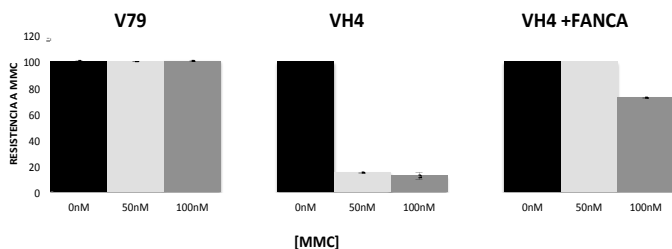
Paralelamente, se realizó el mismo proceso para las células VH4, introduciendo el gen FANCA humano en la línea celular mediante infección por retrovirus y establecimiento de líneas clonales. Posteriormente, se sometió a las células a un experimento de retrotransposición (figura 37B) con la batería de vectores de expresión de LINE1 con selección por blasticidina.

La movilidad del mutante de la reversotranscriptasa de LINE1 que contiene la mutación puntual D702A fue nula en la línea complementada; sin embargo, los niveles de movilidad de los elementos mutantes en la endonucleasa presentaron tasas de movilidad relativamente altas, de 25,5% en el caso de la mutación D205A y 7,7% en la mutación H230A, aunque estos valores están muy alejados de los valores obtenidos en la línea deficiente antes de realizar la complementación, que fueron del 62,5% y el 21,5%, respectivamente (figura 37B).

Para comprobar que realmente la introducción del gen FANCA era responsable también de la reversión del fenotipo, y capaz de eliminar la sensibilidad a agentes intercalantes del DNA, se sometieron las tres líneas celulares a un ensayo de resistencia a MMC (figura 37A). La línea parental, V79, no presentó sensibilidad alguna a la MMC, mientras que la línea deficiente VH4 por el contrario presentó una sensibilidad extrema al agente. La línea VH4 complementada con el gen humano FANCA recuperó la resistencia a MMC a una dosis de 50nM,

mientras que la sensibilidad a dosis de 100nM presentó una resistencia de aproximadamente un 70% con respecto a las células sin tratar, es decir, mostraba una recuperación parcial para las dosis más altas de este agente intercalante del DNA (figura 37A). Estos datos de complementación y la recuperación de la resistencia a MMC en la línea VH4 complementada nos indican que realmente el gen FANCA es el responsable del fenotipo observado.

A)



B)

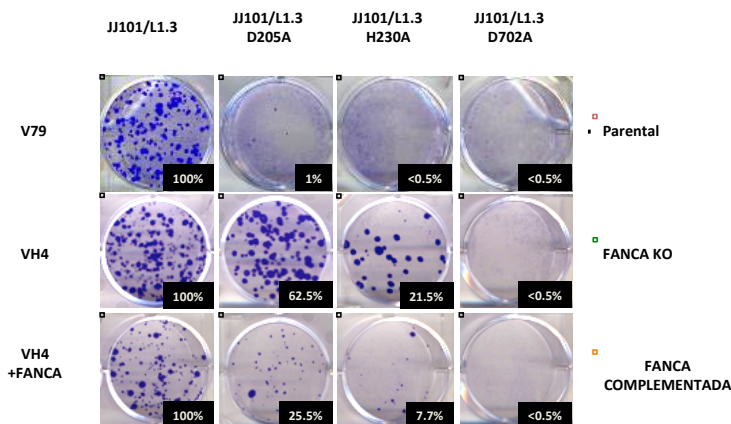


Figura 37. Complementación de la línea VH4 con el gen humano FANCA. A) Ensayo de resistencia a Mitomicina C. La línea complementada con FANCA muestra una corrección total de la sensibilidad a MMC a una dosis de 50nM y parcial (alrededor del 60%) a una dosis de 100nM. B) Ensayo de retrotransposición. La complementación de la línea VH4 con el gen humano de FANCA es capaz de corregir en gran medida las diferencias en los niveles de movilidad de los mutantes en el dominio endonucleasa de L1, disminuyendo hasta 25,5% en D205A y 7,7% en H230A (datos corregidos según eficiencia de transfección).

5.3.- Ensayos de retrotransposición en fibroblastos de pacientes de Anemia de Fanconi.

Posteriormente, tras comprobar la existencia de una desregulación de la movilidad por la vía independiente de endonucleasa en las líneas celulares de hámster, decidimos realizar experimentos de retrotransposición en fibroblastos inmortalizados provenientes de pacientes de anemia de Fanconi (figura 38), con el fin de esclarecer el papel de la ruta FA en la movilidad por esta vía, así como establecer las bases moleculares de la relación de la enfermedad con la movilidad del elemento.

5.3.1.- Ensayos de retrotransposición en fibroblastos de pacientes deficientes en FANCG o FANCD1/BRCA2.

Primero, para asegurar que la ausencia de movilidad de los mutantes en el dominio endonucleasa en las células CHO deficientes en FANCG o BRCA2 no se debe a causas genéticas ajenas a la FA propias de las líneas CHO, se emplearon dos líneas de fibroblastos humanos provenientes de pacientes, concretamente VU-423 y GM16-635, deficientes en BRCA2 y FANCG, respectivamente. La movilidad de los mutantes en el dominio endonucleasa no presentó niveles significativos en ninguno de los dos casos, en comparación con el elemento salvaje, corroborando los resultados obtenidos en las líneas CHO (figura 38). Por tanto, las células de pacientes humanos deficientes en FANCG

o FANC-D1 se comportan como células salvajes, y no se observan niveles de movilidad ENi significativos.

5.3.2.- Ensayos de retrotransposición en fibroblastos de pacientes deficientes en FANCI o FANCD2.

Por otra parte, para aclarar si la movilidad del retroelemento mediante la vía independiente de la endonucleasa se ve afectada realmente por deficiencias en la vía de FA, utilizamos las líneas celulares de fibroblastos humanos PD20F y ANKN, deficientes en los dos efectores principales de la ruta, FANCD2 y FANCI, respectivamente.

En el caso de las células ANKN, deficientes en FANCI, los niveles de movilidad de ambos mutantes del dominio endonucleasa se encontró aumentada, alcanzando un 42% de movilidad en el caso del mutante D205A y hasta un 18% en la mutación más severa H230A (figura 38).

Del mismo modo, las células PD20F, deficientes en FANCD2, presentaron unas tasas de movilidad elevadas de los elementos mutantes en el dominio endonucleasa, con unos valores del 23% en el caso del mutante D205A y un 8% en el caso de H230A (figura 38).

Por tanto, las células de pacientes humanos deficientes en FANC-D2 y FANC-I se caracterizan por albergar niveles elevados de retrotransposición por la vía ENi.

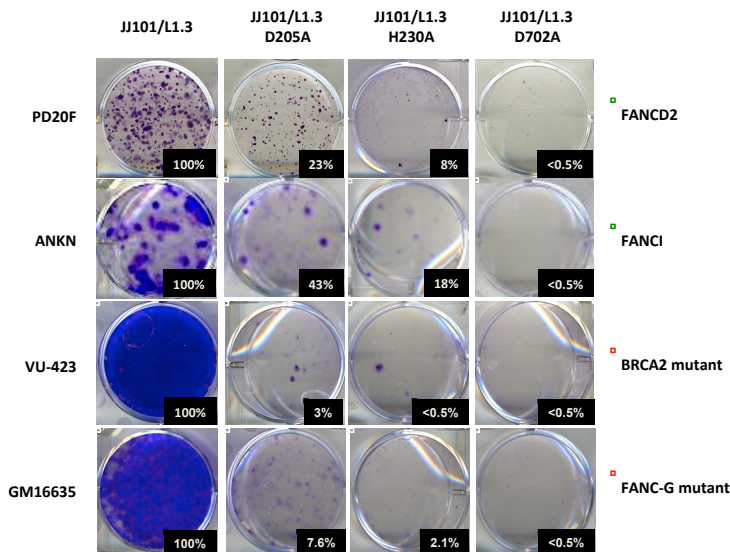


Figura 38. Ensayo de retrotransposición en fibroblastos humanos derivados de pacientes de Anemia de Fanconi. Tras realizar el ensayo, se confirmó el bajo nivel de retrotransposición por la vía ENi en las líneas deficientes en FANCD1/BRCA2 (VU-423) y FANCG (GM16635), con tan sólo un 3% y 7,6% de actividad del mutante D205A, respectivamente. La actividad del mutante H230A fue inferior al 5% en ambos casos (<0,5% y 2,1%, respectivamente). Por el contrario, las células deficientes para los efectores principales de la ruta FA FANCD2 (PD20F) y FANCI (ANKN), mostraron una elevada actividad de los mutantes D205A, con un 23% y un 43% respectivamente, y de los mutantes H230A, con un 8% en PD20F y un 18% EN ANKN (datos corregidos según eficiencia de transfección).

5.4.- Ensayos de retrotransposición en fibroblastos de pacientes de Anemia de Fanconi complementados.

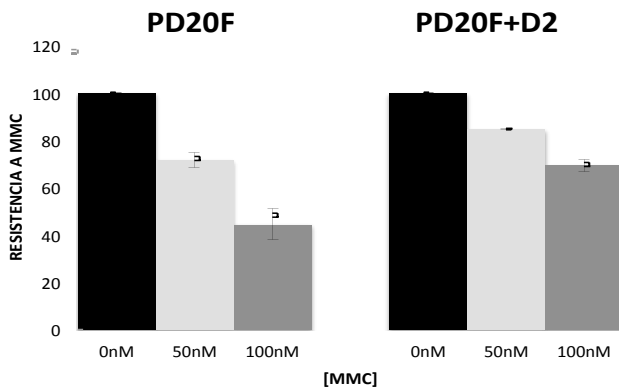
Para comprobar que la movilidad aumentada del elemento mutante en el dominio endonucleasa se debía a la deficiencia en el gen FA afectado, y no a otras causas, se emplearon las líneas ANKN y PD20F complementadas con el gen FA correspondiente. Para demostrar la complementación de las líneas se comprobó el rescate del fenotipo FA mediante un ensayo a resistencia a MMC (figuras 39A y 40A). A diferencia de las líneas celulares de hámster, al tratarse de células obtenidas de pacientes, no se dispone de una línea celular parental que no presente el defecto FA, ya que las células del paciente contienen la mutación original.

5.4.1.- Ensayos de retrotransposición en fibroblastos PD20F complementados con el gen FANCD2.

En el caso de células humanas, en las que no hay una línea parental "sana", para poder demostrar que la línea original del paciente ha sido complementada, seleccionamos el ensayo de resistencia/sensibilidad a MMC. Tras la complementación, la resistencia a MMC de la línea PD20F+FANCD2 (PDF20F+D2) aumentó de manera notable, alcanzando más de un 80% de resistencia a dosis de MMC de 50nM y hasta un 70% de resistencia frente a una dosis de MMC de 100nM, frente a un 70% y un 40%, respectivamente, de la línea original sin complementar (figura 39A).

Una vez determinada de forma experimental la complementación de la línea PD20F con el gen FANCD2, se realizó el ensayo de retrotransposición en la línea complementada. Mientras que los niveles de retrotransposición de los mutantes D205A y H230A en la línea deficiente fueron del 23% y 8%, respectivamente, la movilidad relativa de ambos mutantes descendió hasta niveles indetectables (<0,5%) en la línea complementada PD20F+FANCD2 (figura 39B).

A)



B)

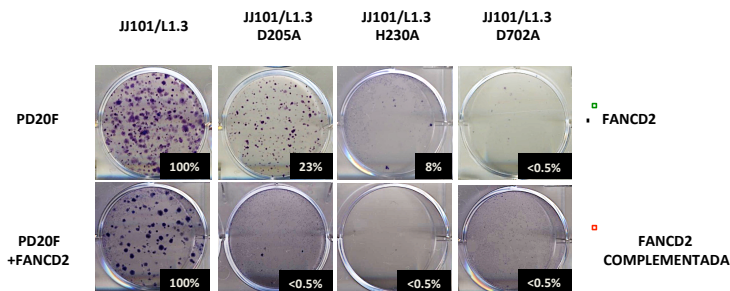


Figura 39. Complementación de los fibroblastos de Anemia de Fanconi PD20F con el gen FANCD2.

A) Ensayo de resistencia a Mitomicina C. La línea complementada con FANCD2 muestra una corrección parcial de la sensibilidad a MMC, aumentando la resistencia del 50% hasta el 72% a una dosis de 100nM. **B) Ensayo de retrotransposición.** La complementación de la línea PD20F con el gen FANCD2 es capaz de corregir por completo los niveles de movilidad de los mutantes en el dominio endonucleasa de L1, disminuyendo hasta niveles casi indetectables en ambos mutantes (<0,5%, datos corregidos según eficiencia de transfección).

5.4.2.- Ensayos de retrotransposición en fibroblastos ANKN complementados con el gen FANCI.

Para comprobar que los cambios en la movilidad de los mutantes de L1 son debidos a la deficiencia en FANCI, se realizó la complementación de la línea ANKN con el gen humano FANCI mediante el uso de un vector viral.

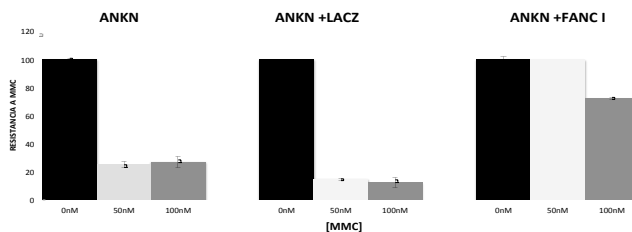
Para demostrar que la adquisición de resistencia a MMC y la disminución de la actividad de los mutantes en la endonucleasa de L1 se debe a la corrección del defecto de la ruta FA, y no al uso de los vectores virales empleados para realizar la complementación *per se*, se generó en paralelo una línea infectada con un vector viral incluyendo el gen *lacZ* como control.

Mediante el ensayo a resistencia a MMC, se comprobó que la línea ANKN complementada con FANCI (ANKN+FANCI), presentaba resistencia total a dosis de 50nM de MMC y valores ligeramente superiores al 70% ante dosis de 100nM. En cambio, la línea parental ANKN deficiente en FANCI, así como la línea control ANKN+LacZ, presentaron valores de extrema sensibilidad ante dosis de 50nM o 100nM de MMC (figura 40A).

Asimismo, tras comprobar la complementación de la línea ANKN con FANCI, se realizó un ensayo de retrotransposición en la línea complementada y en la línea control ANKN+LacZ. Mientras que los niveles de retrotransposición de los mutantes D205A y H230A en la línea control ANKN+LacZ fueron del 46,7% y 18,2%, respectivamente (figura 40B), muy similares a los niveles encontrados en la línea sin complementar

(43% y 18%, respectivamente), la movilidad relativa en la línea complementada ANKN+FANCI descendió hasta un 15,8% en el caso del mutante D205A, y se observó una ligera reducción, de hasta el 12,7%, en el mutante H230A (figura 40B).

A)



B)

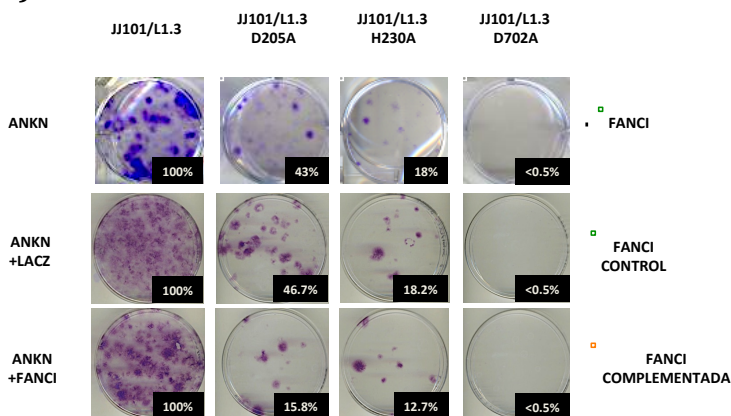


Figura 40. Complementación de los fibroblastos de Anemia de Fanconi ANKN con el gen FANCI. A) Ensayo de resistencia a Mitomicina C. La línea control ANKN+LacZ no mostró corrección de la sensibilidad a MMC al compararla con la línea ANKN, mientras que la línea complementada con FANCI muestra una corrección parcial de la sensibilidad a MMC, aumentando la resistencia del 25% hasta el 72% a una dosis de 100nM de MMC. **B) Ensayo de retrotransposición.** La complementación de la línea ANKN con el gen FANCI corrige los niveles de movilidad de los mutantes en el dominio endonucleasa de L1, disminuyendo la movilidad del mutante D205A del 43-46% en la línea afectada y la línea control hasta un 15,8% en la línea complementada, y la movilidad del mutante H230A desde un 18% hasta un 12,7% (datos corregidos según eficiencia de transfección).

5.5.- Recuperación y caracterización de eventos de retrotransposición.

Las inserciones originadas por la vía de retrotransposición independiente de la endonucleasa o vía de retrotransposición ENi, se encontraron previamente al estudiar células deficientes en la vía de reparación del DNA por la unión de extremos no homólogos (o NHEJ, de *Non-Homologous End Joining*). Estas inserciones originadas por la vía ENi fueron caracterizadas por los autores que las descubrieron (Gilbert et al, 2005; Morrish et al, 2002; Morrish & García-Pérez et al, 2007), y se describió la aparición de determinadas características que no se habían encontrado previamente en las inserciones de L1s salvajes.

Así, las características de las inserciones producidas por elementos salvajes mediante la vía clásica, descritas también en el apartado 2 de introducción, son:

- Truncación en 5' de la secuencia del elemento.
- Presencia de una cola de poliA en la región 3' de la inserción, inmediatamente aguas arriba de la región genómica.
- Aparición de duplicaciones del sitio de inserción (TSDs) de tamaño variable, de 2 a 20 pares de bases.
- La integración se produce en la secuencia de corte consenso reconocida por la endonucleasa de L1 (5'-TTTT/A-3'), o una secuencia similar a ésta.

Por otra parte, las características de las inserciones producidas por mutantes en el dominio endonucleasa mediante la vía ENi, descritas también previamente en el apartado 2 de introducción, son:

- Truncación en 5' de la secuencia del elemento.
- Presencia de la cola de poliA o ausencia de ésta, pudiendo aparecer truncaciones en la región 3' de la secuencia del elemento.
- Alteraciones del sitio de inserción, destacando delecciones del sitio diana asociadas a la inserción.
- La integración se produce en sitios aleatorios en secuencias atípicas, diferentes a la consenso.
- Las inserciones pueden aparecer acompañadas de pequeñas secuencias que, aparentemente, derivan de cDNAs endógenos.

Con el fin de caracterizar las inserciones originadas por la vía ENi en las células deficientes en FA, así como para determinar si éstas se ajustan a las características previamente descritas, se realizaron experimentos de retrotransposición con el casete de recuperación de inserciones (*recovery*). Además de caracterizar inserciones producidas por la vía ENi en las líneas celulares que presentaron movilidad de los mutantes de endonucleasa (VH4, CLV5B, ANKN y PD20F), se caracterizaron inserciones de elementos salvajes en todas las líneas celulares de FA (VH4, CLV5B, ANKN, PD20F, GM16635 y VU423).

5.5.1.- Eventos de retrotransposición recuperados en células de hámster.

Cuando se finalizó el estudio, se caracterizaron completamente un total de 22 inserciones de líneas de hámster deficientes en FA (tabla 8), 11 inserciones producidas por el elemento salvaje (vía TPRT clásica) y 11 inserciones de mutantes D205A producidas por la vía independiente de la endonucleasa (ENi), provenientes de las líneas CLV5B y VH4, deficientes en FANCC y FANCA, respectivamente. Las características encontradas al analizar estas inserciones se resumen en la tabla 9.

Las 22 inserciones caracterizadas poseían las características indicativas de eventos de retrotransposición *de novo*, es decir, presentaban una cola de poliA, el gen de resistencia a neomicina sin intrón y el origen de replicación bacteriano.

5.5.1.1.- Inserciones recuperadas de elementos mutantes en el dominio endonucleasa (D205A)

Al analizar las características de las inserciones obtenidas de elementos mutantes en el dominio endonucleasa, pudimos apreciar la aparición de las duplicaciones del sitio diana o TSDs (*Target Site Duplications*) características de inserciones de LINE-1 (2-20pb aproximadamente) en 2 de los 11 casos, y de acuerdo con la observación de que las inserciones producidas por esta vía se pueden asociar a alteraciones del sitio de integración, se encontraron 2 casos de deleciones en el sitio de inserción, con distinta longitud, una en CLV5B y otra en VH4, de 4 y 126 pares de bases, respectivamente.

Sorprendentemente, las 7 inserciones restantes (2 de CLV5B y 5 de VH4) presentaron una característica no descrita hasta la fecha para las inserciones de L1, la presencia de TSDs de una longitud inusualmente larga, concretamente comprendidas entre 45 y 257 pares de bases (tabla 8), muy superiores a las 2-25 pares de bases esperadas.

5.5.1.2.- Inserciones recuperadas de elementos salvajes (WT)

Aún más sorprendente fue comprobar que, de las 11 inserciones obtenidas de los elementos salvajes (WT), tan sólo 6 presentaban las TSDs características; de los cinco casos restantes, dos presentaron deleciones cortas del sitio de inserción (de 2 y 4pb), y en los tres últimos casos se comprobó de nuevo la aparición de TSDs inusualmente largas, llegando a las 818 pares de bases en una de ellas (tabla 8).

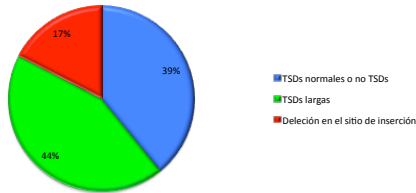
Se encontraron también, entre estas inserciones obtenidas de elementos salvajes, dos casos de eventos que presentaban *"twin priming"*, caracterizados por presentar un fragmento de secuencia de L1 en posición invertida con respecto al resto de la inserción y la cola de poliA; una de las dos inserciones apareció asociada a una deleción de 4pb del sitio de inserción, y la otra presentaba una TSD normal de 13pb.

Nº de inserción	Línea parental	Sublínea	Enzima	Tamaño de la TSD o delección	Longitud de poliA	Truncación de L1 (o Truncación/Inversión)	Localización genómica
13	CLV5B	WT2	HIND3	238bp	A64	5777	gb AFTD01095272.1
54	CLV5B	WT8	HIND3	DEL 4bp	A81	5697	gb AFTD01094172.1
19	CLV5B	ENM5	SSP1	119bp	A46	En gen NEO	gb AFTD01061190.1
55	CLV5B	ENM3	SSP1	DEL 4bp	A57	5711	gb AFTD01038924.1
90	CLV5B	EN2	HIND3	79bp	A40	5640	gb AFTD01017779.1
57	VH4	WT13	HIND3	818bp	A95	5526	gb AFTD01027876.1
58	VH4	WT14	HIND3	15bp	A57	5431	gb AFTD01006972.1
60	VH4	WT8	HIND3	18bp	A40	En gen NEO	gb AFTD01010549.1
61	VH4	WT1	HIND3	102bp	A65	5641	gb AFTD01027694.1
64	VH4	WT9	SSP1	14bp	A104	5799	gb AFTD01012164.1
65	VH4	WT11	SSP1	DEL 2bp	A69	T5300 INV5075-5330	gb AFTD01050786.1
94	VH4	WT2	BGL2	6bp	A60	5859	gb AFTD01108074.1
96	VH4	WT4	BGL2	13bp	A77	T4930 INV4081-4944	gb AFTD01046921.1
97	VH4	WT6	BGL2	BLUNT	A70	5647	gb AFTD01081900.1
66	VH4	EN1	HIND3	DEL 126bp	A47-G3-A21	5855	gb AFTD01099863.1
68	VH4	EN3	HIND3	9bp	A51	5642	gb AFTD01109475.1
70	VH4	EN6	HIND3	70bp	A39	5833	gb AFTD01059582.1
71	VH4	EN14	HIND3	45bp	A37	5928	gb AFTD01138028.1
73	VH4	EN8	HIND3	257bp	A83	5517	gb AFTD01075553.1
76	VH4	EN13	HIND3	80bp	A109	4736	gb AFTD01021298.1
100	VH4	EN2	BGL2	232bp	A46	5427	gb AFTD01016371.1
101	VH4	EN4	BGL2	16bp	A116	5366	gb AFTD01003757.1

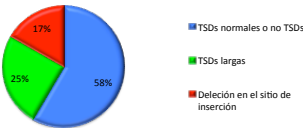
Tabla 8. Inserciones recuperadas en las líneas celulares de Hámster.

De izquierda a derecha se indican: número de inserción o clon recuperado; línea parental; sublínea de la que se obtuvo la inserción; enzima utilizada para recuperarla; longitud de la TSD o delección, indicando en azul las TSDs de longitud normal (1-25pb o “blunt” para 0bp), en verde las TSDs largas (más de 25pb) y en rojo las delecciones en el sitio de inserción; longitud de la cola de poliA; posición nucleotídica de L1 en la que se truncó la retrotransposición del elemento, en morado y anotadas como “INV” se indican las inserciones que presentaron truncación e inversión (“twin priming”), indicándose en estos casos además las posiciones de inicio y fin de la secuencia invertida del L1; por último, se indica la localización genómica en la que se produjo la inserción, referida a GenBank, ya que el genoma de Hámster no se había ensamblado por completo en el momento en el que se realizó este estudio.

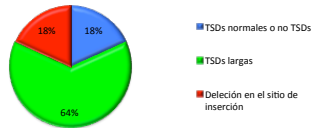
Inserciones en células de Hámster



Inserciones de elementos WT en células de Hámster



Inserciones por la vía ENi en células de Hámster



..

Inserciones de las líneas	TSDs normales o no TSD	TSDs largas	Delección en el sitio de inserción	Número total
CL-V5B WT	0	1	1	2
CL-V5B ENi	0	2	1	3
VH4 WT	6	2	1	9
VH4 ENi	2	5	1	8

Tabla 9. Características de las inserciones recuperadas en las líneas celulares de Hámster.

Se indica la línea parental de procedencia y el elemento del que se obtuvo la inserción: WT para los eventos que se obtuvieron de un elemento salvaje (*Wild type*) y ENi para los que proceden de un evento de retrotransposición independiente de endonucleasa (mutantes D205A del dominio endonucleasa en todos los casos).

5.5.2.- Eventos de retrotransposición recuperados en fibroblastos humanos.

Se caracterizaron un total de 38 inserciones obtenidas en las líneas procedentes de fibroblastos deficientes en la ruta FA. Todas ellas presentaron las características indicativas de eventos de retrotransposición *de novo*, como son la presencia de una cola de poliA y del origen de replicación bacteriano y el gen NEO sin el intrón. Las características de las inserciones se resumen en la tabla 12.

5.5.2.1.- Inserciones recuperadas de elementos mutantes en el dominio endonucleasa (D205A)

Debido a la enorme dificultad para obtener sublíneas que acomodaran la movilización de mutantes en el dominio endonucleasa en las líneas parentales VU-423 y GM-16635, ya que presentan un nivel muy bajo o casi inexistente de retrotransposición por la vía ENi, sólo se obtuvieron inserciones de elementos salvajes para estas líneas (tabla 11), concretamente ocho de cada una de ellas, las veintidós inserciones restantes caracterizadas pertenecen a las líneas PD20F y ANKN, de las cuales diez se obtuvieron de mutantes en el dominio endonucleasa y doce de elementos WT (tabla 10).

De las inserciones que se produjeron por la vía ENi, cuatro presentaron las TSDs típicas de longitud normal (0-25pb), mientras que se encontraron cuatro inserciones asociadas a deleciones en el sitio de inserción y, por último, las

dos inserciones restantes presentaron de nuevo TSDs largas, de 27 y 47pb.

Nº de inserción	Línea parental	Sublínea	Enzima	Tamaño de la TSD o deleción	Longitud de poliA	Truncación de L1	Localización cromosómica
1	PD20F	WT2	HIND3	BLUNT	A99	En gen NEO	15q21.1
3	PD20F	WT6	HIND3	BLUNT	A74	5922	18q12.2
5	PD20F	WT6	HIND3	38bp	A135	5435	8p11.22
39	PD20F	WT7	HIND3	23bp	A44	5276	4q26
40	PD20F	WT7	HIND3	53bp	A26	5940	5p12
44	PD20F	WT11	HIND3	10bp	A82	5445	14q24.3
7	PD20F	EN2B	HIND3	DEL 3bp	A91	5667	3q23
47	PD20F	EN30B	HIND3	27bp	A73	5477	20q13.2
79	PD20F	EN22B	BGL2	47bp	A35	5969	Yp11.2
80	PD20F	EN13B	BGL2	DEL 25bp	A83	3321	5q11.2
21	ANKN	WT7	HIND3	DEL 6bp	A159	4557	4q31.3
22	ANKN	WT2	HIND3	DEL 3bp	A21	5671	2p25.1
24	ANKN	WTM2	SSP1	11bp	A86	T5254 INV1736-1778	6q21
84	ANKN	WT2B	BGL2	16bp	A37	5553	13q31.1
85	ANKN	WT14	BGL2	DEL 5-10bp	A68	5574	3p14.3
162	ANKN	WT16C	HIND3	DEL 7bp	A27	3853	4q26
30	ANKN	EN4	HIND3	4bp	A60	4179	11q21
31	ANKN	EN1	SSP1	1bp	A37	5092	4p15.2
33	ANKN	ENM2	SSP1	BLUNT	A27	5352	2q23.3
34	ANKN	EN9B	HIND3	DEL 4928bp	A32	5834	11p11.2
35	ANKN	EN17B	HIND3	1bp	A69	5455	12q14.3
166	ANKN	EN11D	HIND3	DEL 291bp	A7	5507	12q14.3

Tabla 10. Inserciones recuperadas en las líneas celulares humanas PD20F y ANKN. De izquierda a derecha se indican: número de inserción o clon recuperado; línea parental; sublínea de la que se obtuvo la inserción; enzima utilizada para recuperarla; longitud de la TSD o deleción, indicando en azul las TSDs de longitud normal (1-25pb o “blunt” para 0bp), en verde las TSDs largas (más de 25pb) y en rojo las deleciones en el sitio de inserción; longitud de la cola de poliA, posición nucleotídica de L1 en la que se truncó la retrotransposición del elemento, en azul y anotadas como “INV” se indica una inserción que presentó *twin priming*, indicándose las posiciones de inicio y fin de la secuencia invertida del L1; por último, se indica la localización cromosómica en la que se produjo la inserción.

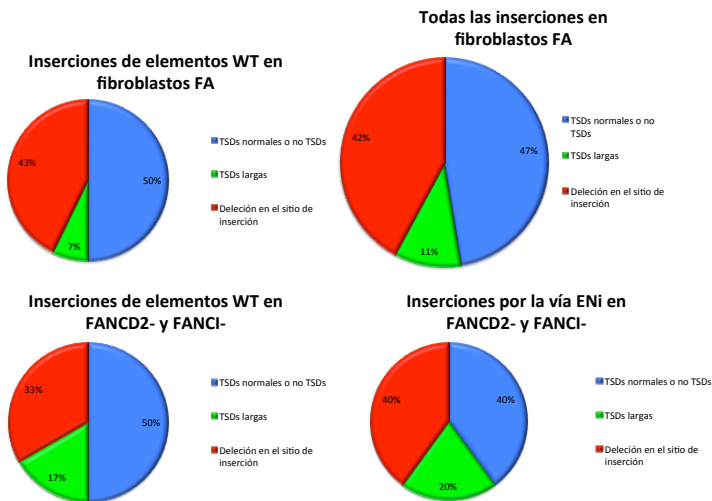
5.5.2.2.- Inserciones recuperadas de elementos salvajes (WT)

Al analizar las inserciones obtenidas de elementos salvajes en las líneas deficientes en FANCD2 (PD20F) y FANCI (ANKN), se volvió a comprobar la escasa representación de inserciones normales, y tan sólo un 50% de los eventos totales (6 eventos de los 12) presentaron TSDs de longitud normal. 4 inserciones estaban asociadas a deleciones del sitio de inserción y, de nuevo, 2 inserciones presentaron TSDs largas, concretamente de 38 y 53pb (tabla 10).

La mayor sorpresa apareció al analizar las inserciones de elementos salvajes en las líneas deficientes en BRCA2 (VU-423) y FANCG (GM-16635). Se esperaba encontrar TSDs de longitud normal, pero sólo aparecieron en ocho de las dieciséis inserciones, las ocho restantes presentaron, al igual que otras líneas FA, deleciones en el sitio de inserción con un tamaño variable, desde tan sólo 1 nucleótido hasta una deleción de tamaño significativo, de más de 400kb, en una sublínea de GM-16635. En estas líneas no se encontró la presencia de TSDs largas en ninguno de los casos hasta el momento de presentar este estudio (tabla 11).

Nº de inserción	Línea parental	Sublínea	Enzima	Tamaño de la TSD o delección	Longitud de poliA	Truncación de L1	Localización cromosómica
103	VU-423	WT1	HIND3	DEL 9bp	A99	5966	4q32.3
106	VU-423	WT6	HIND3	DEL 19bp	A113	5483	12q23.1
108	VU-423	WT8	HIND3	11bp	A79	5711	5q31.3
109	VU-423	WT9	HIND3	DEL 7bp	A94	5868	6q15
111	VU-423	WT12	HIND3	8bp	A64	5524	20p11.21
134	VU-423	WT2	SSP1	6bp	A70	4930	15q15.2
135	VU-423	WT3	SSP1	14bp	A32	4573	12q13.13
136	VU-423	WT3	SSP1	DEL 3760	A39	5445	16p13.11
114	GM-16635	WT13	HIND3	4bp	A28	5911	2q21.1
115	GM-16635	WT8	HIND3	DEL 11bp	A36	5857	12q14.1
117	GM-16635	WT8	HIND3	12bp	A84	T5613 INV3688-3859	11q14.2
149	GM-16635	WT9	SSP1	15bp	A72	5009	20p12.3
150	GM-16635	WT9	SSP1	DEL 1bp	A66	5851	11p15.1
151	GM-16635	WT6	SSP1	3bp	A58	5531	1p34.2
153	GM-16635	WT5	SSP1	DEL 404098bp	A36	5622	8p22
154	GM-16635	WT5	SSP1	DEL 14172bp	A35	5332	17q24.1-17q23.3

Tabla 11. Inserciones recuperadas en las líneas celulares humanas VU-423 y GM-16635. De izquierda a derecha se indican: número de inserción o clon recuperado; línea parental; sublínea de la que se obtuvo la inserción; enzima utilizada para recuperarla; longitud de la TSD o delección, indicando en azul las TSDs de longitud normal (1-25pb) y en rojo las delecciones en el sitio de inserción; longitud de la cola de poliA, posición nucleotídica de L1 en la que se truncó la retrotransposición del elemento, en azul y anotadas como “INV” se indica una inserción que presentó “*twin priming*”, indicándose las posiciones de inicio y fin de la secuencia invertida del L1; por último, se indica la localización cromosómica en la que se produjo la inserción.



∴

Inserciones de las líneas	TSDs normales o no TSD	TSDs largas	Delección en el sitio de inserción	TOTAL
PD20F WT	4	2	0	6
PD20F ENi	0	2	2	4
ANKN WT	2	0	4	6
ANKN ENi	4	0	2	6
VU-423 WT	4	0	4	8
GM-16635 WT	4	0	4	8

Tabla 12. Características de las inserciones recuperadas en las líneas celulares derivadas de fibroblastos humanos. Se indica la línea parental de procedencia y el elemento del que se obtuvo la inserción: WT para los eventos que se obtuvieron de un elemento salvaje (“*Wild type*”) y ENi para los que proceden de un evento de retrotransposición independiente de endonucleasa (mutantes D205A del dominio endonucleasa en todos los casos).

5.5.3.- Análisis de las inserciones anómalas recuperadas en líneas celulares deficientes en la ruta FA.

5.5.3.1.- Inserciones con TSDs inusualmente largas.

Un 64% de las inserciones producidas por la vía ENi en las líneas de hámster presentaron duplicaciones del sitio de inserción (TSDs) inusualmente largas, llegando incluso hasta 818 pares de bases.

En las líneas humanas el porcentaje de inserciones producidas por la vía ENi que presentaron estas TSDs largas se situó en un 20%, con una longitud máxima de 53 pares de bases.

El porcentaje de inserciones mediadas por elementos salvajes presentando TSDs largas fue del 25% en líneas de hámster y del 17% en las líneas humanas.

En la figura 41 se muestra, como ejemplo representativo de este tipo de eventos, la inserción número 100, obtenida de la línea VH4, con una TSD de 232pb.

Nº 100 (L1 D205A) 232bp TSD

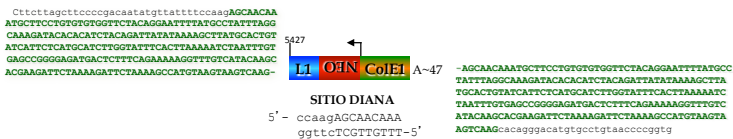


Figura 41. Ejemplo de inserción con TSDs inusualmente largas. Representación gráfica de la caracterización de la inserción número 100, con una TSD de 232 pb. Se muestran los nucleótidos que flanquean la inserción, indicándose el punto de truncación de la secuencia de L1, la

longitud de la cola de poliA y el sitio diana (o pre-integración). En verde se indican los nucleótidos duplicados a ambos lados de la inserción. Al tratarse de genoma de hámster no se puede determinar su ubicación cromosómica.

5.5.3.2.- Inserciones asociadas a deleciones en el sitio de inserción.

Las deleciones cortas, de unos pocos nucleótidos en el sitio de inserción, se encontraron en un gran número de las inserciones, apareciendo en todas las líneas deficientes en FA analizadas; asimismo, se detectaron varias inserciones asociadas a deleciones de tamaño considerable en el sitio de inserción, desde 100 pares de bases hasta miles de pares de bases.

En total, un 18% y un 40% de las inserciones producidas por la vía ENi en las líneas de hámster y humanas, respectivamente, presentaron algún tipo de deleción del sitio de integración asociada, siendo la más notable de casi 5Kb en una de las inserciones recuperadas de la línea humana ANKN.

En los eventos producidos por elementos salvajes, la proporción de deleciones asociadas al sitio de inserción en células deficientes en FA se situó en el 17% en las células de hámster y un 43% en las líneas humanas, encontrándose incluso una deleción de más de 400Kb en la línea VU423.

En la figura 42 se muestra, como ejemplo representativo de este tipo de eventos, la inserción número 154, obtenida de la línea GM16635, con una pérdida de más de 14Kb de la secuencia genómica en el sitio de inserción. Todos estos datos ponen de manifiesto que la actividad del retroelemento L1

puede contribuir a la alta inestabilidad genómica que presentan estas células humanas.

N° 154 (L1 WT) Delección de 14.172bp

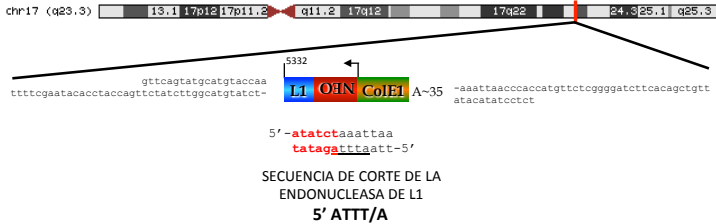


Figura 42. Ejemplo de Inserción asociada a una delección genómica. Representación gráfica de la caracterización de la inserción recuperada número 154. Se muestra la ubicación genómica de la inserción con una imagen procedente del *Genome Browser*, así como los nucleótidos que flanquean la inserción, indicándose el punto de truncación de la secuencia de L1 y la longitud de la cola de poliA. En la parte inferior se indica la secuencia de reconocimiento de la endonucleasa de L1 en el sitio diana, indicándose en rojo el inicio de la secuencia genómica perdida a causa de la inserción.

5.5.3.3.- Inserciones con secuencias de relleno.

Las inserciones acompañadas de pequeñas secuencias nucleotídicas se han descrito previamente como parte normal de eventos de retrotransposición producidos por la vía ENi (Morrish et al, 2002) y ocasionalmente en eventos producidos por la vía clásica (TPRT) (Beck et al, 2010); no obstante, normalmente estas adiciones de material genético no exceden de unos pocos nucleótidos. En el caso de las inserciones recuperadas en anemia de Fanconi, pudimos detectar 4 inserciones que presentaban la adición de secuencias inusualmente largas de DNA en

el extremo 5', en uno de los casos (inserción número 39), como algo curioso, cabe destacar además que esta secuencia presentaba fragmentos cortos reestructurados y repetidos en bloques (figura 43).

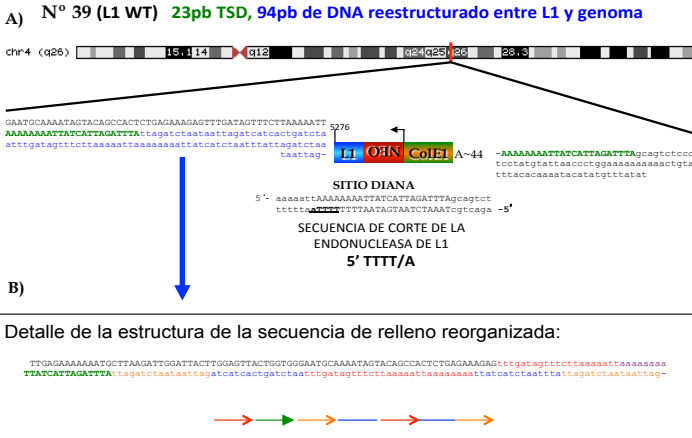


Figura 43. Ejemplo de una inserción con secuencia de relleno. La figura muestra la inserción recuperada número 39. **A) Representación gráfica de la caracterización del evento de retrotransposición.** Se muestra la ubicación genómica de la inserción con una imagen procedente del *Genome Browser*, así como los nucleótidos que flanquean la inserción, indicándose el punto de truncación de la secuencia de L1 y la longitud de la cola de poliA. Los nucleótidos en verde indican las TSDs y los nucleótidos en azul la secuencia de relleno ajena al elemento y a la región genómica. En la parte inferior se indica la secuencia de reconocimiento de la endonucleasa de L1 en el sitio diana. **B) Detalle de la reorganización del sitio genómico donde se produjo la inserción.** Las letras de cada color indican las secuencias repetidas, indicando en verde la TSD, en la parte inferior, las flechas de color indican la posición relativa de las secuencias repetidas.

5.6.- Ensayos de retrotransposición en células tratadas con RNAs de interferencia.

Para poder recapitular de una forma aún más fidedigna los resultados obtenidos en los experimentos de retrotransposición en ausencia de algunas de las proteínas de la ruta FA, recurrimos al empleo de siRNAs. Esta aproximación presenta dos ventajas importantes: la primera es que el fondo genético es idéntico en todas las células tratadas, y la segunda es que nos permite eliminar de una forma específica la expresión de una proteína concreta de la ruta FA. Así, podríamos transfectar las células con siRNAs frente a los distintos miembros de la ruta FA y realizaríamos los ensayos de retrotransposición pertinentes.

Lo primero que hicimos fue optimizar el periodo de espera desde que se realiza la cotransfección de los siRNAs y los vectores de expresión de L1, hasta que se añade el antibiótico blasticidina al medio para realizar la selección; como resultado, se puede concluir que la selección con blasticidina se puede empezar hasta 48h después de la transfección de los vectores de expresión de L1, sin perder por ello un número notable de colonias al final del experimento (Figura 42).

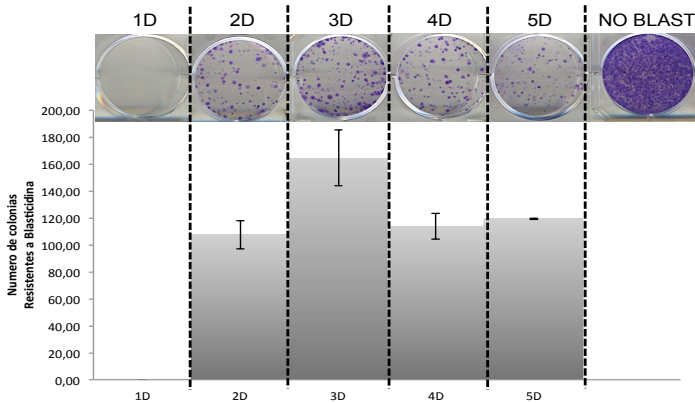
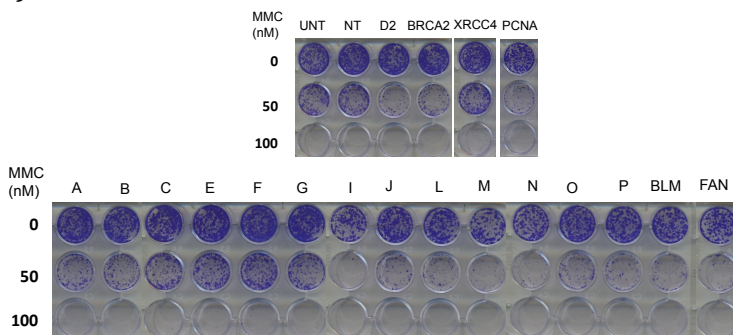


Figura 42. Ensayo de retrotransposición con selección por blasticidina. Se muestra el número de colonias resultantes tras 8 días de selección con blasticidina, añadida a las células a las 24, 48, 72, 96 o 120 horas (días 1 al 5) tras la transfección del vector JJ101/L1.3. En la parte superior, a la derecha, se muestra también una tinción de las colonias resultantes de células transfectadas con el vector JJ101/L1.3 sin seleccionar (no blast).

Asimismo, para demostrar que la inhibición específica de las proteínas FA se produce al usar los siRNAs y que se reproduce el fenotipo de la enfermedad, se realizaron experimentos de resistencia a MMC en las células tratadas con los siRNAs (figura 43). En concreto se utilizaron siRNAs para los 15 miembros de la FA (A, B, C, D1, D2, E, F, G, I, J, L, M, N, O y P), así como para disminuir otras proteínas que guardan una estrecha relación con la ruta FA, concretamente la nucleasa FAN1, la helicasa BLM, y la plataforma deslizante del DNA PCNA. Como controles se incluyeron células sin tratar, células tratadas con siRNA sin diana específica (*Non Targeting* o NT) y tratadas con el siRNA frente a XRCC4, que no afecta a la ruta FA.

A)



B)

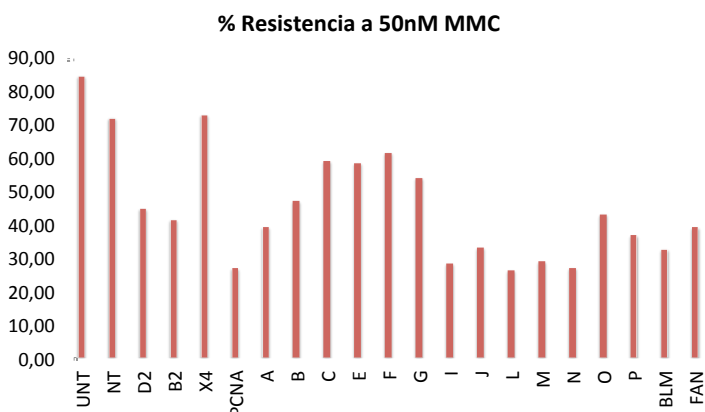


Figura 43. Ensayo de resistencia a mitomicina C con siRNAs en células HCT. A) La imagen muestra la tinción con cristal violeta del ensayo de resistencia a diferentes dosis de mitomicina C (0, 50 y 100nM). **B)** La gráfica muestra el porcentaje de resistencia a una dosis de mitomicina C de 50nM, obtenido al comparar por espectrofotometría la intensidad de tinción de las líneas tratadas con el siRNA indicado, tras someterlas a una dosis de 50nM de mitomicina C durante 72 horas, con respecto al valor obtenido con el mismo siRNA sin tratar con mitomicina C (0nM MMC).

Tras comprobar que los siRNAs inhiben la ruta FA, y que se puede empezar la selección con blasticidina 72 horas tras la transfección, realizamos los experimentos en diversas condiciones y diferentes líneas celulares, con el fin de comprobar el efecto de la inhibición de las proteínas de la vía FA sobre la vía de retrotransposición ENi; desafortunadamente, en ninguno de los casos se consiguió obtener diferencias significativas en la movilidad de los mutantes en el dominio endonucleasa (datos no mostrados), ni siquiera en el control positivo XRCC4, cuya ausencia permite una gran movilidad del elemento L1 por la vía ENi, pero que no afecta a la retrotransposición del elemento salvaje (Morrish et al, 2002).

No obstante, y dada la existencia de inserciones anómalas producidas por la vía clásica en todas las líneas examinadas, y con el fin de determinar el posible efecto de las proteínas FA en la movilidad del elemento, realizamos el experimento con elementos salvajes y los siRNAs, en las líneas celulares HeLa, U2OS y HCT (figura 44).

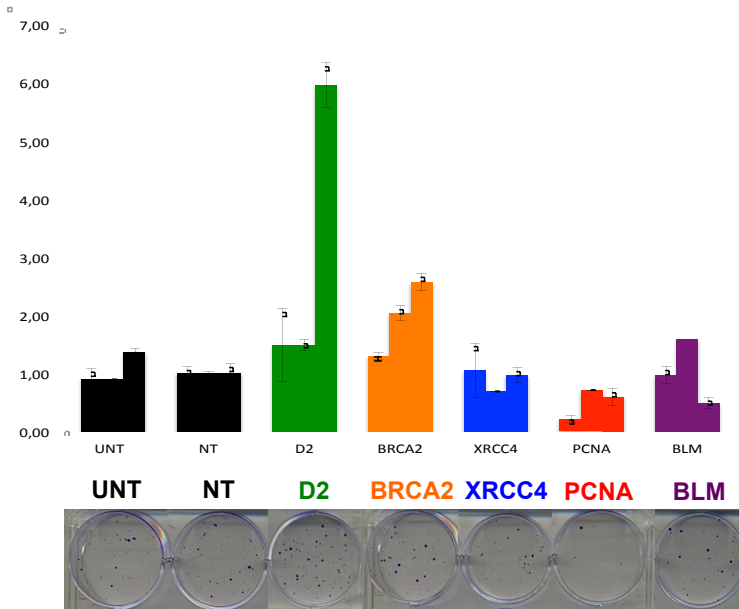


Figura 44. Resultados del ensayo de retrotransposición de L1 WT ante deficiencias en diferentes rutas de reparación. Comparación de los niveles de retrotransposición de elementos L1 salvajes ante deficiencias en los diferentes componentes del complejo FA, aunque la ausencia de uno de ellos suele provocar fallos en los demás, se indica el proceso de reparación principalmente afectado: coordinación de anemia de Fanconi (FA) en verde; FA y recombinación homóloga (HR) en naranja; unión de extremos no homólogos en azul; síntesis translesión en rojo; recombinación homóloga en morado. Cada barra representa el resultado en cada línea celular, las líneas celulares de izquierda a derecha son: HCT, U2OS y HeLa. En la parte inferior se muestra una imagen de las colonias obtenidas en la línea HCT (columnas de la izquierda de cada grupo).

Aunque los resultados presentaron una alta variabilidad en muchos casos, se pudo observar un incremento en los niveles de retrotransposición de los elementos salvajes ante deficiencias en FANCD2 y BRCA2, en comparación con las células tratadas con los siRNAs control, en todas las líneas analizadas, alcanzando un incremento de hasta 6 y 2,5 veces, respectivamente, en los niveles de retrotransposición en células HeLa. No se observó una variación significativa en la movilidad del elemento frente a la inhibición de XRCC4, ni una variación consistente ante inhibición de la helicasa BLM. Por otra parte, sorprendentemente, la inhibición de PCNA redujo drásticamente los niveles de retrotransposición de los elementos salvajes en todas las líneas celulares estudiadas, dándose la reducción más drástica en la línea HCT, donde el nivel de retrotransposición cayó hasta el 21% (reducción de 5 veces) en comparación con las células tratadas con el siRNA sin diana (NT).

Por otra parte, también analizamos el resto de miembros de la ruta FA y la nucleasa asociada FAN1, los resultados obtenidos se muestran en la figura 46. A pesar de la alta variabilidad encontrada en algunos de los miembros, la inhibición de la mayoría de las proteínas FA causó un incremento en la retrotransposición de L1, a excepción de FANCF, FANCG y FANCM, cuya inhibición causó una disminución notable de los niveles de retrotransposición. Todos los resultados fueron comparados con un control de expresión de blastomicina para evitar posibles sesgos causados por diferencias en la viabilidad celular.

Todos estos resultados confirman que la actividad de L1 se encuentra desregulada ante

deficiencias en la reparación por la vía FA, y sugieren una implicación importante de la actividad del elemento en la reparación del DNA en las células deficientes en los mecanismos de reparación y replicación.

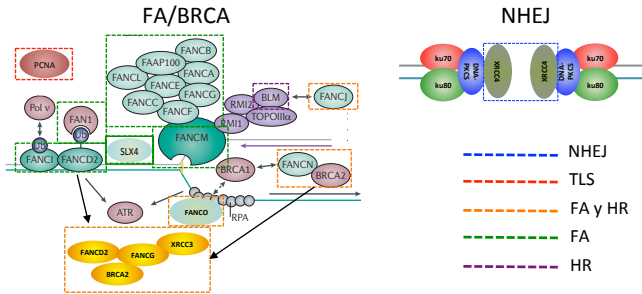
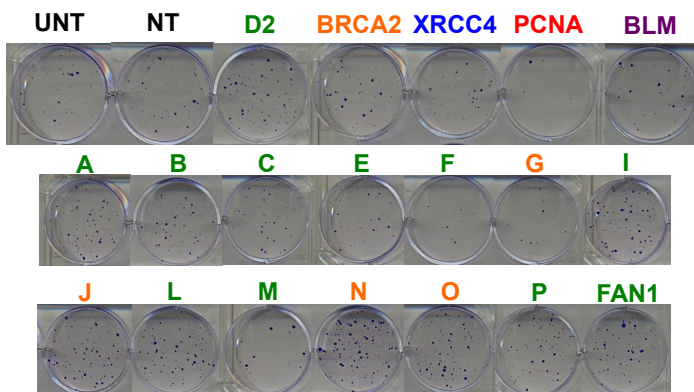


Figura 45. Esquema de la ruta FA y proteínas accesorias. Los recuadros punteados indican los complejos inhibidos experimentalmente mediante siRNA. Los diferentes colores indican el proceso afectado principalmente: FA, coordinación de anemia de Fanconi (FA) en verde; FA y HR, anemia de fanconi y recombinación homóloga en naranja; NHEJ, unión de extremos no homólogos en azul; TLS, síntesis translesión en rojo; HR, recombinación homóloga en morado.

A)



B)

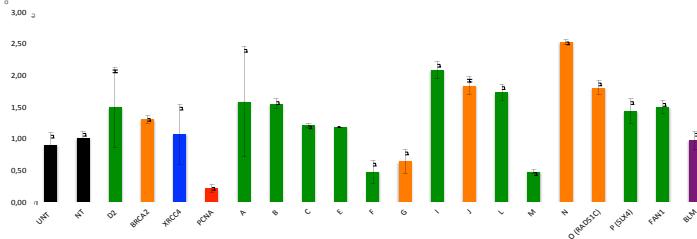


Figura 46. Ensayo de retrotransposición de L1 WT con siRNAs en células HCT. A) Resultados del ensayo de retrotransposición en células HCT. **B)** Variación de los niveles de retrotransposición de elementos L1 ante inhibición de las proteínas FA, se indica la variación en la retrotransposición con respecto al siRNA control sin diana, establecido como 1, las barras indican la desviación estándar.

DISCUSIÓN

6.- DISCUSIÓN.

El retroelemento L1 es un agente biológico mutagénico inherente al ser humano, dado que su actividad ha mostrado ser causa de la aparición de enfermedades genéticas y cáncer en numerosas ocasiones en los últimos años (Kazazian et al., 1988; Miki et al., 1992; Narita et al., 1993; Schwahn et al., 1998; Yoshida et al., 1998; Meischl et al., 2000; Takai et al., 2000; Iskow et al., 2010; Lee et al., Science 2012; Soylom et al., Genome Res 2012; Shukla et al., 2013; revisado en Beck et al., 2011).

No obstante, las causas subyacentes a este efecto mutagénico de L1 han permanecido elusivas, y nuestros resultados muestran, por primera vez, una asociación del retroelemento con un síndrome de inestabilidad genética como es la enfermedad humana Anemia de Fanconi. Esta relación queda demostrada por la alta incidencia de inserciones anómalas independientemente de la vía de retrotransposición, y el aumento de la movilidad basal de los elementos, tanto salvajes como mutantes en su dominio endonucleasa.

Así pues, aunque no se conoce con exactitud cuántos elementos L1 mutantes en el dominio endonucleasa pueden habitar actualmente el genoma humano, sí que se han encontrado ejemplos de movilización de estos elementos en él (Sen et al., 2007; Beck et al., 2010); aunque este hecho no representa una amenaza en condiciones normales, dado que su movilidad es extremadamente baja (Seleme et al., PNAS 2006), puede convertirse en una de las causas principales de la predisposición a tumores que se observa en la patología anemia de Fanconi (revisado en Alter, 2003), y probablemente, de otros síndromes donde la reparación del DNA se encuentra alterada.

6.1.- Actividad desregulada de L1 en la Anemia de Fanconi.

En cuanto a la relevancia clínica de nuestros datos, debemos recordar que la prevalencia de FANCA y FANCC constituye cerca del 80% de todos los casos de FA, seguidos por FANCG que representa un 9-10% de todos los casos y FANCD2 y BRCA2, que representan en torno al 2-3% cada uno, mientras que FANCI se encuentra mutado en un 1-2% de los pacientes de FA (revisado en Moldovan & D'Andrea, 2009). Así, las mutaciones en alguno de los miembros estudiados en esta Tesis mediante ensayos celulares (A, C, G, D1, D2 e I), son responsables de cerca del 90% de los casos de FA presentes en la clínica. También cabe destacar la asociación del gen BRCA2 con el cáncer de mama (Howlett et al, 2002).

Nuestros resultados demuestran que las células deficientes en los genes FANCA, FANCC, FANCD2 y FANCI son capaces de soportar eventos de retrotransposición de L1 producidos por la vía independiente de su endonucleasa (ENi), y que dicha vía de movilidad no se produce de manera habitual en las células deficientes en los genes BRCA2 y FANCG.

6.1.1.- La retrotransposición de L1 por la vía ENi está aumentada en líneas celulares de hámster y fibroblastos de pacientes con defectos en la vía de reparación FA.

El porcentaje de actividad de los elementos mutantes en el dominio endonucleasa (ENi) denominados D205A, en células de Hámster respecto al de los elementos salvajes fue del 62.5% en células

deficientes en FANCA, frente a un 1% en la línea parental y un 25.5% en la línea complementada con FANCA. En el caso de FANCC la movilidad del mutante D205A fue del 25%, frente a un 1% en la línea parental y un 4.8% en la línea complementada con FANCC. En fibroblastos deficientes en FANCD2, su actividad se situó en un 23% frente a menos de un 0.5% en la línea complementada con FANCD2. Finalmente, en los fibroblastos deficientes en FANCI la movilidad del mutante D205A fue del 43% respecto de los elementos salvajes, mientras que en los fibroblastos complementados con FANCI se redujo hasta el 15.8%. Por tanto, los valores de movilidad de L1 mediante la vía ENi hallados en estas líneas celulares, deficientes en la ruta de reparación FA, se sitúan muy por encima de los niveles relativos que se obtienen normalmente para este mutante, que generalmente se sitúan entre el 0.5 y el 5% (Moran et al.,1996; Morrish et al, 2002).

6.1.2.- La complementación de líneas deficientes en miembros de FA revierte el aumento de movilidad de los elementos L1 mutantes en distintos miembros de la vía de FA.

Tras la complementación de las líneas deficientes con el gen de la ruta FA correspondiente, se produce un notable incremento de la resistencia a Mitomicina C (MMC) en las células complementadas, a la vez que el nivel de movilidad de los mutantes D205A de L1 cae drásticamente; estos cambios no se observan al introducir el gen control lacZ, demostrando por tanto que es la introducción del gen FA lo que corrige la sensibilidad extrema a la MMC y

reduce los niveles de movilidad de los mutantes de L1.

Por otra parte, la movilidad de L1 por la vía ENi, aun a pesar de verse drásticamente reducida al corregir la deficiencia en la ruta FA, se sigue produciendo con cierta frecuencia en algunos casos, aunque es importante destacar que la complementación del fenotipo FA se realizó generando líneas multiclonales con un vector viral; por tanto, es posible que el transgen que introducimos para corregir la deficiencia sea silenciado en algunas de las células complementadas, en función del contexto genómico de la región en la que se integró. Por otra parte, en el caso de FANCA y FANCC, la corrección parcial de los niveles de movilidad de los mutantes de L1 en las células de Hámster podría deberse a que la complementación se realizó con los genes humanos, por diferencias en los procesos de reparación entre especies; de hecho, esta corrección parcial también se observa a la hora de comprobar la sensibilidad a MMC de las células complementadas, en las que no se restituye completamente la resistencia de la línea parental salvaje, a pesar de apreciarse un incremento notable con respecto a la línea mutante sin complementar.

6.1.3.- Las inserciones de L1 en líneas deficientes en la ruta FA presentan características anómalas asociadas a procesos de inestabilidad genómica.

Asimismo, un exhaustivo análisis de las inserciones que hemos obtenido revela varias características anómalas en una amplia proporción de las inserciones analizadas (superior al 50%). Destacamos dos, la presencia de largas duplicaciones en el sitio de inserción, y la presencia de deleciones. La presencia de duplicaciones del sitio de inserción inusualmente largas se ha detectado en un 44% de las inserciones analizadas en líneas de Hámster y en un 11% de las recuperadas en fibroblastos humanos. Por otra parte, la presencia de deleciones en el sitio de inserción pudo observarse en aproximadamente el 17% de las inserciones recuperadas de líneas de Hámster y un 42% de las inserciones humanas. Estos resultados demuestran que la actividad de L1 es capaz de promover la inestabilidad genética en las células deficientes en FA, aunque resulta imposible determinar con exactitud el alcance de la implicación de la actividad de L1 en este proceso de inestabilidad genómica característico de la anemia de Fanconi.

Además, en el análisis de inserciones de elementos salvajes, hemos encontrado una gran proporción de ellas asociadas a modificaciones anómalas del sitio de inserción (alrededor del 50%), un comportamiento no descrito hasta la fecha para elementos salvajes. Nuestros resultados demuestran que estas anomalías están asociadas al proceso de retrotransposición que tiene lugar en las células deficientes en miembros de FA, independientemente de si ésta se produce por la vía clásica (TPRT) o la vía ENi, y revelan que existe una asociación entre estas

inserciones anómalas y los defectos moleculares de la patología.

Aunque la presencia de deleciones o duplicaciones largas del sitio de inserción (TSDs), producidas por la actividad del elemento L1, se han descrito previamente (Gilbert et al., 2002), estos eventos son muy poco frecuentes en condiciones normales. Cabe destacar que las deleciones asociadas al sitio de inserción sí se han observado con cierta frecuencia, en eventos de retrotransposición producidos por la vía ENi, ante deficiencias en la vía de reparación por unión de extremos no homólogos o NHEJ (Morrish et al, 2002 & 2007).

No obstante, resulta muy difícil discernir si el hecho de que la actividad de L1 se encuentre desregulada en las células deficientes en la ruta FA es tan sólo una consecuencia directa de los problemas que se producen en los procesos de reparación o si, por otra parte, es un componente principal de la etiogénesis propia de la patología, a modo de causa de inestabilidad genética y/o de la aparición de tumores. Lo más interesante es que ambas explicaciones no son mutuamente excluyentes, por lo que ambos sucesos podrían estar ocurriendo simultáneamente, como parecen indicar los datos incluidos en esta Tesis; no obstante, no se puede descartar que el papel que juega L1 en la patología, a nivel molecular, no sea tan importante como otros defectos en la reparación o la replicación intrínsecos de la enfermedad.

6.1.4.- El silenciamiento de genes de la ruta FA y genes asociados altera los niveles de retrotransposición de L1.

Sorprendentemente, no hemos podido detectar la movilidad de L1 por la vía ENi al utilizar siRNAs para inhibir las proteínas de reparación estudiadas, incluyendo XRCC4, cuya ausencia causa una elevada movilidad de L1 por la vía ENi (Morrish et al 2002 & 2007); es importante destacar que los siRNAs, aunque causan una reducción severa de los niveles de la proteína a la que inhiben, no generan una pérdida total de la misma como ocurre en líneas mutantes aisladas de pacientes o del modelo de hámster. Este resultado negativo podría reflejar que, tanto en NHEJ como FA y HR, la célula puede evitar que se produzca la retrotransposición por la vía ENi incluso a niveles muy bajos de reparación funcional.

Sin embargo, en los experimentos con RNAs de interferencia y elementos L1 salvajes, se observa una variación de la movilidad de L1 al transfectar las células con siRNAs frente a la mayoría de las proteínas FANCA, FANCB, FANCC, BRCA2, FANCD2, FANCE, FANCI, FANCI, FANCL, FANCN, FANCO, FANCP y FAN1; además, hemos podido detectar una reducción notable de la movilidad de los elementos salvajes tras la inhibición de FANCG, FANCF y FANCM. Hemos obtenido resultados similares para PCNA (*Proliferating Cell Nuclear Antigen*), aunque es importante destacar que esta proteína está directamente relacionada con la replicación (Prelich & Stillman, 1988) y la maquinaria de la

recombinación homóloga (Li et al., 2009; Li et al., 2013), y a pesar de su estrecha relación con la ruta FA (Howlett et al, 2009; Moldovan et al, 2007), no forma parte directa de ella como tal.

Al inhibir la expresión de XRCC4 mediante siRNAs, no se observa ninguna variación significativa en los niveles de retrotransposición de elementos salvajes, como se esperaba por lo demostrado en estudios anteriores (Morrish et al, 2002 & 2007).

Todos estos datos parecen indicar que la bajada de los niveles de las proteínas FANC es suficiente para que la retrotransposición por la vía clásica se produzca con más frecuencia. Esto podría deberse a que se requiera más tiempo para reparar las lesiones en el DNA, y esa demora aumente la probabilidad de que se produzca una integración exitosa de los retroelementos salvajes.

6.2.- Inestabilidad genómica causada por la actividad de L1 en la Anemia de Fanconi.

Para el funcionamiento de una célula, así como para un desarrollo correcto de los individuos en última instancia, es fundamental que el genoma se mantenga bien estructurado y estable; sin embargo, a lo largo del ciclo celular, las células tienen que enfrentarse a diferentes accidentes o agentes, externos e internos, que pueden dañar o alterar la estructura de su material genético, lo que requiere la intervención coordinada de la maquinaria de reparación del DNA (revisado en Deans & West, 2011).

La actividad de L1 puede ser causante de alteraciones cromosómicas, tales como deleciones en el sitio de inserción (Gilbert et al., 2002; Symer et al., 2002), así como traslocaciones al recombinarse dos secuencias homólogas, ya sean de L1 o Alu (Lehrman et al., 1987, Hu et al., 1991) que se encuentran en localizaciones distintas del genoma. Por tanto, la célula necesariamente debe ejercer un control estricto sobre la movilidad de estos elementos con el fin de preservar su estabilidad genómica.

Una de las principales características de la disrupción de la vía FA es el aumento significativo de aberraciones cromosómicas que se detectan al exponer las células a agentes que causan daños en el DNA, debido a la incapacidad de reparar de una manera coordinada dichos daños, lo que se traduce en una elevada inestabilidad genómica (revisado en Moldovan & D'Andrea, 2009).

Los resultados de esta Tesis demuestran que, al menos ante deficiencias en los genes FANCA, FANCC, FANCD2 y FANCI, existe un incremento notable de la movilidad de los elementos L1 por la vía independiente de la endonucleasa (ENi), lo que sugiere que las ribonucleopartículas de L1 son capaces de aprovechar las roturas no reparadas del DNA para movilizarse.

En más del 50% de las inserciones de L1 analizadas en las células deficientes en la ruta de reparación de FA, se observa la aparición de características asociadas a procesos de inestabilidad genómica, como duplicaciones del sitio de inserción (TSDs) anormalmente largas, una alta prevalencia de deleciones y la adición de largos fragmentos de DNA de relleno. Hemos encontrado estas características en la región del genoma donde se produjo la integración independientemente de si las inserciones del elemento L1 se produjeron por la vía ENi o por la TPRT. Por tanto, nuestros resultados indican que las alteraciones genómicas producidas por la actividad de L1 se encuentran sobrerrepresentadas en las células FA, apuntando a una implicación de la actividad de L1 en la inestabilidad genómica encontrada en estas células y, por otra parte, a que la intervención de la vía FA juega un papel crucial en la resolución de las estructuras intermediarias de la retrotransposición. Este papel podría ser el de evitar que los eventos de retrotransposición se completen con éxito, o que se produzcan alteraciones del locus genómico debido a la actividad del retrotransposón.

6.3.- Vía de retrotransposición independiente de endonucleasa de L1 ante deficiencias en la reparación del DNA.

Como ya hemos comentado, los distintos tipos de lesiones que se producen en el DNA obligan a la célula a dar una respuesta adecuada, siendo necesaria una coordinación correcta de las diferentes vías de reparación del DNA (revisado en Bogliolo et al., 2002). A veces, cuando alguno de estos mecanismos no funciona correctamente, otros pueden ofrecer una solución alternativa pero, generalmente, con una mayor predisposición al error, que puede provocar la aparición de un amplio abanico de mutaciones o aberraciones cromosómicas (revisado en Bogliolo et al., 2002). Así, una deficiencia en el funcionamiento de algún mecanismo de reparación, especialmente en el caso de la ruta coordinadora FA, puede conllevar un uso excesivo de mecanismos alternativos, y un aumento en la probabilidad de introducir errores en la reparación. El resultado, un amplio espectro en la frecuencia y gravedad de alteraciones genéticas presente en la sintomatología de las enfermedades causadas por fallos en la reparación del DNA (revisado en Thompson & Hinz, 2009).

Es precisamente en el uso de mecanismos alternativos de reparación del DNA donde puede darse la participación accidental de LINE-1, como ya demostraron otros autores ante deficiencias en la vía de unión de extremos no homólogos (Morrish et al., 2002 & 2007) y, más recientemente, en la reparación de roturas de doble cadena (Onozawa et al., 2014). Los resultados presentados aquí indican que L1 es

capaz de aprovechar los fallos en la coordinación de la reparación y replicación del DNA para movilizarse, incluso cuando su endonucleasa no puede participar en el proceso de retrotransposición. Pero la pregunta más evidente ante este fenómeno es: ¿cuáles son las consecuencias de la movilidad de L1 por la vía ENi?

Hasta ahora hemos considerado el aspecto negativo de las inserciones del retroelemento, pero en las células deficientes en la vía FA podrían tener distintos efectos:

- Por un lado, si la lesión sin reparar que aprovecha L1 para integrarse estaba en una región cuya expresión no es esencial para la función que desempeña la célula, ni altera su proliferación, esta reparación alternativa podría permitir a la célula sobrevivir y dividirse. En este caso, la actividad de LINE-1 resultaría beneficiosa tanto para la célula como para el individuo.

- Por otro lado, la lesión puede producirse en una región importante para la función celular y/o el control de la proliferación (Shukla et al., 2013; Helman et al., 2014). Una inserción de LINE-1 en dichas regiones perjudica o beneficia a la célula como ente individual, pero generalmente perjudica a la población celular normal y por tanto al individuo si la célula no es eliminada del organismo.

Es difícil imaginar la intervención de L1 como beneficiosa para el individuo, ya que no deja de ser un elemento insercional y por tanto un agente mutagénico (revisado en Babushok & Kazazian, 2007; Beck et al., 2011). Ahora bien, en una célula que no es capaz de coordinar la maquinaria adecuada, el uso accidental de la maquinaria de L1 a modo de mecanismo de reparación alternativo (Onozawa et al.,

2014), puede resultar beneficioso si es capaz de reparar una lesión que, de otro modo, no sería eliminada.

6.4.- Reparación, replicación del DNA y actividad de L1.

La actividad de los elementos móviles constituye la mayor fuente de variabilidad genética conocida (revisado en Cordaux & Batzer 2009), y esto significa que su persistencia en los genomas está favorecida por la acción de la selección natural. La mutación es tanto admisible como necesaria a nivel de evolución; sin embargo, a nivel de individuo, la coexistencia de elemento móvil y genoma requiere forzosamente mantener un equilibrio entre mutación y reparación (revisado en Babushok & Kazazian, 2007), ya que la generación de nueva información a partir de los elementos móviles supone un escenario de posibilidades muy variadas:

- Puede que la nueva información genética no tenga ninguna repercusión a nivel evolutivo en la especie.
- Puede que, en el momento de producirse la inserción o tras cambios puntuales sucesivos, esta información dé lugar a la aparición de nuevas funciones o genes (Sayah et al, 2004)
- Puede aportar inestabilidad genómica en la región génica en la que se produce, con la consiguiente susceptibilidad a la aparición de algún tipo de enfermedad genética asociada a un mal funcionamiento de la región donde se produjo el evento (Shukla et al, 2013; Helman et al., 2014).

Todas estas consideraciones parecen carecer de importancia en un individuo, pero si aplicamos el mismo concepto considerando a la célula como individuo y al ser humano como entorno, donde conviven varias poblaciones celulares, desde el punto de vista de una célula, si es capaz de adquirir una ventaja proliferativa sobre el resto y evitar ser eliminada del organismo, independientemente de los efectos sobre el ser humano, irá aumentando su representación en la población celular, convirtiéndose en la “especie dominante”. En muchas ocasiones, este fenómeno suele ir acompañado de una pérdida o disminución de la capacidad de dichas células para desempeñar correctamente su función, lo que acaba teniendo consecuencias negativas sobre el organismo, como son los tumores o el cáncer (revisado en Burrell et al., 2013).

Por tanto, para desempeñar correctamente las funciones complejas, es importante que el organismo sea capaz de mantener bajo control tanto la proliferación como la función de las diferentes poblaciones celulares que lo constituyen. La actividad del retroelemento L1 añade sin duda una capa más de mutaciones con pronóstico incierto.

Así, la combinación de los efectos acumulados de las mutaciones por agentes externos y las causadas por la actividad desregulada de L1 puede ser devastadora para los pacientes de anemia de Fanconi; de hecho, recientemente, se ha demostrado que, una alta tasa de retrotransposición se asocia con una alta tasa de reordenaciones cromosómicas y mutaciones somáticas en cáncer (Helman et al., 2014). Nuestros resultados apuntan a que esto es precisamente lo que ocurre de forma frecuente en el caso de los pacientes

de anemia de Fanconi, y posiblemente, en otros síndromes causados por fallos en la reparación del DNA.

Uno de los ejemplos de enfermedades genéticas donde esta actividad descontrolada del L1 podría ser de vital importancia es el cáncer de mama, donde una mutación en una sola copia de los genes BRCA1, BRCA2, FANCI, FANCD1 o FANCD2, predispone a su aparición (revisado recientemente en Kotteman & Smogorzewska, 2013). Curiosamente, nuestros datos revelan que una inhibición parcial de cualquiera de estas proteínas mediante siRNAs, supone un incremento notable de la actividad de los elementos L1 salvajes. De hecho, la influencia de BRCA1 sobre la retrotransposición ya ha sido descrita por otros autores, junto con la influencia de ATM (Coufal et al., 2011).

6.4.1.- Maquinaria de LINE-1 y la reparación del DNA en FA.

La proteína ORF2p posee dos claros dominios:

- Una endonucleasa de tipo APE (apurínica/apirimidínica (Martin et al., 1995; Martin et al., 1996; Feng et al., 1996), que es capaz de realizar cortes en secuencias similares a la secuencia consenso, lo que abre la oportunidad de provocar una mella o una rotura de doble cadena en puntos cercanos a una lesión.
- La reversotranscriptasa de L1, que se comporta no sólo como una RT, sino que además es capaz de actuar utilizando DNA como molde como una polimerasa celular (Mathias et al., 1991; Piskareva & Schmatchenko, 2006).

Existen varias posibilidades a la hora de evaluar qué ocurre en las células FA (figura 46), ya que en estas células, las horquillas de replicación que se han bloqueado o colapsado no se consiguen restablecer eficientemente, y por tanto la señalización celular provoca un arresto del ciclo celular y eventualmente la muerte de la misma. En estas células, si la actividad dual de L1 puede estar ayudando de alguna manera al restablecimiento de una horquilla dañada, aunque sea de manera mutagénica, puede resultar muy útil para continuar con el ciclo celular.

La actuación desorganizada y aleatoria de proteínas de diferentes rutas de reparación junto con la actividad de L1 explicaría el amplio espectro de características anómalas que se encuentran en las regiones en las que se producen las inserciones de L1 en las células FA. De hecho, los resultados obtenidos en este estudio apuntan a que estas anomalías pueden estar produciéndose al intentar resolver las estructuras creadas durante el proceso de retrotransposición de L1.

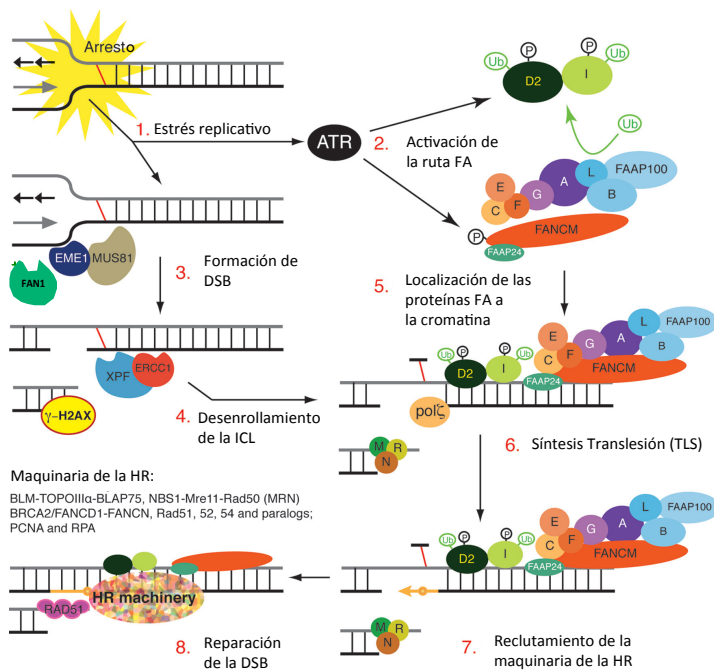


Figura 46. Reparación de cruzeamientos intercatenarios (ICLs) por la vía FA. Los ICLs detienen el avance de la horquilla de replicación, provocando el estrés replicativo que conduce a la parada del ciclo celular (1) y a la activación de ATR y la ruta FA (2), que promueven la formación de un corte en el DNA que crea una rotura de doble cadena (DSB) inicial (3). Esto permite desenrollar la ICL mediante otro corte en el DNA (4), así como la localización del complejo FA en la lesión para ubiquitinar a FANCD2 y FANCI *in situ* (5). Se produce así la llamada de una polimerasa de reparación (junto con PCNA) para realizar la síntesis translesión (6). Posteriormente, se recluta la maquinaria de recombinación homóloga (7), que repara finalmente la rotura de doble cadena formada al principio de la reparación (8), restableciéndose así la horquilla de replicación.

6.4.2.- Recombinación homóloga acoplada a la replicación y actividad de L1 en FA.

Una célula siempre intenta reparar las roturas de doble cadena de la manera más rápida posible, y durante la interfase ésta es la vía de unión de extremos no homólogos o NHEJ. Sin embargo, durante la replicación, la ruta FA inhibe la NHEJ y establece una manera de reparar el daño mediante la coordinación de nucleasas y polimerasas de la recombinación homóloga y de la síntesis translesión (revisado en Deans & West, 2011).

En el procesamiento necesario para la reparación de las roturas es frecuente la resección de los extremos, y la exposición de filamentos de DNA de cadena sencilla (ssDNA) que son rápidamente recubiertos por RAD51 (Sung & Robberson, 1995). La presencia de ssDNA en el genoma es reconocida por el complejo 9-1-1, ATR-ATRIP y BRCA1. ATR recluta al complejo ID de la vía FA, y a su vez reclutan a complejos remodeladores de la cromatina (CRCs), a PCNA, y a RPA para que recubra el ssDNA, desplazando así a RAD51 y por tanto favoreciendo la TLS en contra de la HR. Por otro lado BRCA1 favorece la acumulación de RAD51 aguas abajo en el filamento de cadena sencilla, promoviendo la HR frente a la TLS (revisado en Thompson, 2012).

En las deficiencias de la vía FA, donde este mecanismo no se produce de manera adecuada, cabría esperar que el peso de la reparación quede en manos de nucleasas e invasiones de cadena, propias de la recombinación homóloga. Nuestros resultados demuestran que la movilidad por la vía ENi no se produce ante deficiencias en los genes FANCD1 (BRCA2) ni FANCG, lo que pone de manifiesto que

una deficiencia en la reparación de las roturas de doble cadena del DNA no es suficiente para que los elementos L1 puedan movilizarse por la vía ENi; no obstante, curiosamente estas dos proteínas, no sólo participan en la FA, también interaccionan entre sí y son necesarias para la recombinación homóloga (HR) (Hussain et al., 2003), por lo que las células con mutaciones en BRCA2 o FANCG son también deficientes en la recombinación homóloga (Yamamoto et al., 2003; Nagaraju & Scully, 2007; Wilson et al., 2008)

Así pues, según nuestros datos, los eventos por la vía ENi pueden producirse ante deficiencias en la ruta FA, pero no se producen ante deficiencias de la HR; sin embargo, FANCG forma parte del complejo principal FA, y la recombinación homóloga como tal forma parte de la reparación por la vía FA, así que ¿cómo es posible entonces? La unión de ambas observaciones, aparentemente contrarias, podría ser explicada de dos maneras que no son mutuamente excluyentes:

- Puede que las primeras reacciones de la reparación, que funcionan de manera independiente y previa a la recombinación homóloga (figura 46), sean suficientes para producir unos cambios en la rotura que impidan que L1 pueda acceder a ella, esto consistiría en la acción coordinada de la resección de los extremos y polimerasas de reparación, que podrían dejar la rotura accesible a otra maquinaria como la de NHEJ, que, aunque descoordinada y con una gran predisposición a error, sigue funcional en estas células. Cabría esperar deleciones de las zonas dañadas ante este escenario.

- Por otra parte, cabe la posibilidad de que los mutantes en el dominio endonucleasa de L1 requieran la participación de parte de la maquinaria de la recombinación homóloga para completar su movilización; esto es, incluso aun siendo capaces de iniciar la reversotranscripción de su DNA parasitando una rotura preexistente, al no ser capaces de producir ningún corte adicional en el sitio diana, quedan a merced de la recombinación homóloga para completar exitosamente su integración.

En circunstancias normales, la ruta FA sería capaz de coordinar la síntesis translesión, sobrepasando la lesión, reiniciando así la replicación sin necesidad de que la horquilla de replicación colapse, y evitando por tanto que se produzca una rotura de doble cadena (figura 47). El complejo principal FA, así como FANCD1 y BRCA1 (posiblemente junto con su acompañante BARD1), pueden intervenir para evitar un posible bloqueo estructural que pueda surgir como consecuencia de la lesión.

Sin embargo, ante determinadas circunstancias, como ocurre por ejemplo con los cruzamientos intercatenarios (ICLs), la formación de una rotura de doble cadena es necesaria para la reparación. En este caso la horquilla de replicación colapsa y se requiere la coordinación de la TLS, la HR y la vía de escisión de nucleótidos (NER) para poder restablecer la horquilla de replicación (figura 48, revisado en Wang, 2007).

En las células FA, la incapacidad de coordinar la TLS, posiblemente incluso ante otras lesiones distintas de los ICLs, provoca la formación de una rotura y por tanto el colapso de la horquilla de

replicación (revisado en Wang, 2007), haciendo necesario recuperarla para poder continuar con la replicación (figura 48). En este caso, la célula sólo podría recurrir a la NHEJ o la HR, ambas funcionales; sin embargo, además de los problemas para coordinar la TLS, las células FA también presentan defectos en la coordinación de las distintas vías de reparación (revisado en Wang, 2007), lo que implica un retraso en la ejecución de cualquiera de estas dos alternativas. Esta circunstancia puede estar dando más tiempo al retroelemento para insertarse.

Además, si tenemos en cuenta el mecanismo de movilización de L1, durante la reversotranscripción se produce un híbrido RNA-DNA de L1 (Luan et al., 1993), con lo que si se degrada el RNA, un fragmento de DNA de cadena sencilla quedaría al descubierto, independientemente de que el elemento L1 tenga su endonucleasa funcional o no, dejando a la célula FA las mismas opciones que ante una estructura intermediaria de cadena sencilla, es decir, la NHEJ o la HR.

Así, por un lado, la intervención de la NHEJ y la vía de escisión de nucleótidos, ambas funcionales a pesar de los fallos en la FA (revisado en Niedernhofer et al., 2005), podrían resultar en la aparición de las deleciones encontradas en todas las líneas FA analizadas.

La búsqueda forzada de homología podría provocar un segundo punto de “ruptura” más lejano de lo normal en el DNA, aguas abajo, en la cadena contraria, lo que explicaría tanto los altos niveles de movilidad de los elementos L1 mutantes en su endonucleasa, como la aparición de las largas TSDs que hemos podido detectar en los grupos FANCA,

FANCC, FANCI y FANCD2, que alcanzan hasta las 800 pares de bases.

Consecuentemente con esta observación, en las células deficientes para BRCA2 o FANCG, el escenario es completamente diferente, ya que sólo la primera opción está disponible, destacando que:

- Deficiencias en BRCA2 no provocan fallos en la ubiquitinación de FANCD2, y por lo tanto la localización del complejo ID en la cromatina no está impedida, y las células pueden coordinar correctamente la síntesis translesión, pero no la recombinación homóloga.

- Deficiencias en FANCG sí que provocan fallos en la ubiquitinación y localización de FANCD2 y la localización del complejo ID, además de fallos en la recombinación homóloga.

En ambos casos, como indican nuestros resultados, las inserciones de L1 ocurren mediante la vía de TPRT clásica, es decir, la endonucleasa del elemento se requiere para su movilidad. Cabe destacar que, según los datos obtenidos de los experimentos con RNAs de interferencia, la movilidad de los elementos L1 se encuentra aumentada en células deficientes para BRCA2, presentando un 130% de actividad comparado con el RNA de interferencia control en células HCT116, en las que encontramos el incremento de actividad más modesto, llegando incluso a superar el 250% de actividad con respecto al RNA control en células HeLa y U2OS.

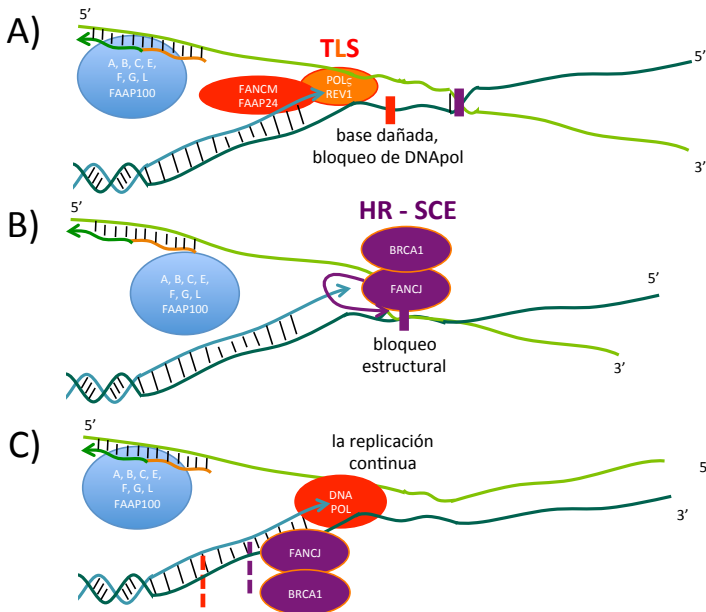


Figura 47. Reinicio de una horquilla de replicación bloqueada. Una de las funciones de la vía FA es desbloquear la replicación de la cadena de DNA, coordinando la TLS y la HR o el intercambio de cromátidas hermanas (SCE). La acción de las proteínas de la vía FA: **(A)** promueve la síntesis translesión sin realizar una rotura de doble cadena, **(B)** cualquier bloqueo estructural o superenrollamiento que pueda impedir la progresión de la horquilla de replicación es resuelto o relajado mediante la participación de la helicasa FANCI asociada a BRCA1, **(C)** permitiendo que continúe la replicación.

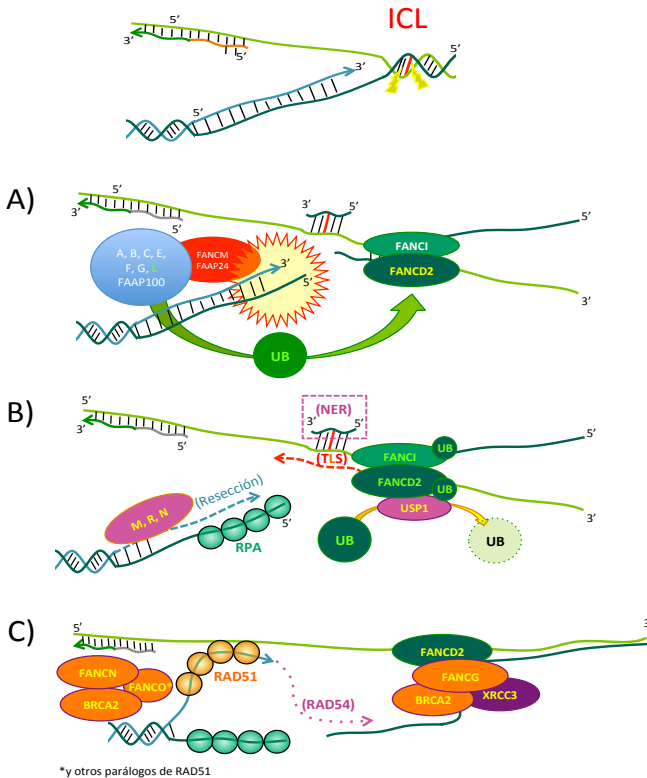


Figura 48. Recuperación de una horquilla de replicación colapsada.

En ocasiones, bien sea por un bloqueo que no se consigue eliminar o a causa de una DSB, la horquilla de replicación colapsa. Durante la reparación de ICLs, eliminar el bloqueo requiere precisamente la formación de una DSB y por tanto el colapso de la horquilla de replicación. Para restablecer una horquilla colapsada se requiere la intervención coordinada de las proteínas de la vía FA: **(A)** primero el complejo principal FA ubiquitina al complejo ID, que coordina la síntesis translesión, **(B)** luego USP1 retira la ubiquitina del complejo ID, permitiendo que FANCG forme el complejo D1-D2-G-X3, **(C)** que recluta a la maquinaria de la recombinación homóloga para poder restablecer eficientemente la horquilla y continuar la replicación.

En el caso de FANCG; sin embargo, aunque la movilidad del elemento por la vía independiente de endonucleasa aparece también impedida, los elementos salvajes no presentan una mayor tasa de movilidad. De hecho, según los resultados con RNAs de interferencia, la deficiencia en FANCG parece disminuir el nivel de movilización de los elementos L1 salvajes, como indica el dato de una actividad relativa del 70% frente a células tratadas con el siRNA control.

Todo esto, unido a que las deficiencias en la vía de NHEJ provocan un aumento tremendo de la movilidad de los mutantes de endonucleasa de L1, apunta a la necesidad de la recombinación homóloga para la correcta resolución de los eventos de retrotransposición producidos por la vía ENi. Además, la presencia de una endonucleasa en el ORF2, podría permitir a L1 integrarse sin requerir la participación de la maquinaria de la recombinación homóloga, en tal caso la célula podría recurrir a mecanismos de reparación alternativos, principalmente la NHEJ, para evitar la pérdida de material genético; si bien, si ésta falla, aún podría recurrir a la NER para eliminar el material genético que no puede reparar, pudiendo dar lugar a las deleciones genómicas y/o truncaciones en 3' del elemento observadas en estudios previos ante deficiencias en la vía NHEJ (Morrish & García-Pérez et al, 2007), pero ausentes en nuestro estudio. Esto se debe probablemente a que las células FA poseen una NHEJ funcional, de hecho, una de las funciones principales propuestas de la vía FA es la inhibición de la NHEJ (Adamo et al., 2010).

A priori, lo que se podría esperar es que la no ubiquitinación de FANCD2 es lo que permite la movilidad de los elementos mutantes para la endonucleasa, pero FANCG es necesario para la ubiquitinación de FANCD2 y su ausencia parece no permitir la movilización de los mutantes de L1. Es muy destacable la participación conjunta de estas dos proteínas en el complejo D1-D2-G-X3. Estos datos, difíciles de interpretar, hicieron que centráramos nuestra atención en este subcomplejo D1-D2-G-X3, cuya formación depende de FANCG y es independiente de la ubiquitinación de FANCD2, pero aún no se conoce con exactitud la función que desempeña (Wilson et al, 2008). Nuestros datos apuntan a un posible papel de este subcomplejo en la integración o resolución exitosa de los eventos de retrotransposición iniciados por los mutantes de L1.

6.4.3.- Replicación del DNA y actividad de L1.

A pesar de la caracterización y el creciente conocimiento de las propiedades y procesividad de la RT del elemento L1 humano (Mathias et al., 1991; Piskareva & Schmatchenko, 2006; Taylor et al., 2013), aun no se ha conseguido determinar con exactitud cómo se produce la síntesis de la cadena complementaria del DNA del elemento, necesaria para completar los eventos de retrotransposición. Puede ser que esta síntesis sea llevada a cabo por la propia RT del elemento, como ocurre con el retrotransposón R2 (Christensen & Eickbush, 2005). Por otra parte, como se demostró, el elemento L1 es capaz de moverse en cualquier momento del ciclo celular (Kubo et al, 2006), lo que sugiere que la RT de L1 es capaz por sí sola de completar todo el proceso

de síntesis de DNA en un evento de retrotransposición *de novo* (figura 49).

Este mecanismo sería suficiente para explicar la aparición de TSDs normales en todos los tipos celulares analizados en este estudio (y en estudios previos); sin embargo, nuestros experimentos presentaron un resultado altamente sorprendente: la inhibición de PCNA, la plataforma deslizante del DNA, mediante siRNAs, redujo drásticamente los niveles de retrotransposición del elemento, disminuyéndolos hasta un 12% con respecto a las líneas tratadas con el siRNA control, o lo que es lo mismo, una reducción de 9 veces en el nivel de retrotransposición normal.

Así pues, debido a que PCNA se expresa de manera mayoritaria durante la fase S (Matthews et al., 1984; revisado en Moldovan et al., 2007), nuestros datos obligaban a considerar que el elemento se moviliza activamente durante la fase S, mientras el DNA se está replicando, y aquí la situación es mucho más compleja por la presencia de la maquinaria de replicación y dos dobles cadenas de DNA.

PCNA es una proteína que funciona durante la replicación, rodeando al DNA. Funciona como una plataforma o andamio para que otras proteínas interaccionen con él, destacando las polimerasas de replicación del DNA (Oku et al., 1998; Zhang et al., 1999; revisado en Moldovan et al., 2007). La mayoría de las proteínas conocidas que interaccionan con PCNA presentan un motivo de unión a esta proteína, conocido como PIP-box (por *PCNA-interacting protein box*) (Warbrick, 1998), generalmente compuesto por la secuencia consenso “**Q-x-x-h-x-x-aa**”, donde “Q” representa Glutamina, “h” representa aminoácidos con cadenas laterales moderadamente hidrofóbicas,

como por ejemplo Leucina, Isoleucina o Metionina (L, I, o M), “a” representa aminoácidos con cadenas laterales fuertemente aromáticas, como por ejemplo Fenilalanina o Tirosina (F o Y), y “x” representa cualquier aminoácido (Warbrick, 1998; Warbrick et al., 1998). Una de las proteínas que contiene este dominio de interacción con PCNA es precisamente FANCD2 (Howlett et al, 2009), de suma importancia, ya que se ha demostrado que las polimerasas translesión pueden interactuar con FANCD2 antes que con PCNA, y mediante los mismos dominios de interacción (Fu et al, 2013). Por otra parte, es especialmente destacable el hecho de que ya se ha demostrado la interacción de PCNA con los transposones de DNA *Pogo*, en *Drosophila*, mediante este dominio de interacción. Este dominio se encuentra también en su homólogo humano *Tigger* (Warbrick et al., 1998), así que la interacción entre PCNA y elementos móviles ya se ha producido previamente en la evolución.

Mediante un estudio de la secuencia aminoacídica de la ORF2p de un elemento L1HS consenso, comprobamos que existe esta secuencia consenso en varios elementos L1 de diferentes especies de mamíferos. En el caso del ser humano la secuencia es “**Q-T-T-I-R-E-Y-Y**”, en las posiciones aminoacídicas 408-415 del ORF2 humano, concretamente en el dominio Z de la proteína, situado entre los dominios endonucleasa y reversotranscriptasa, y cuya función aún no se conoce. Es importante destacar que, durante el progreso de nuestro estudio, otro grupo que estudiaba interacciones de las proteínas ORF1 y ORF2 mediante coinmunoprecipitación, encontró esta

interacción (Taylor et al., 2013), por lo que, a pesar de no conocerse su funcionalidad, la interacción de la proteína ORF2 con PCNA ya ha sido demostrada (Taylor et al., 2013).

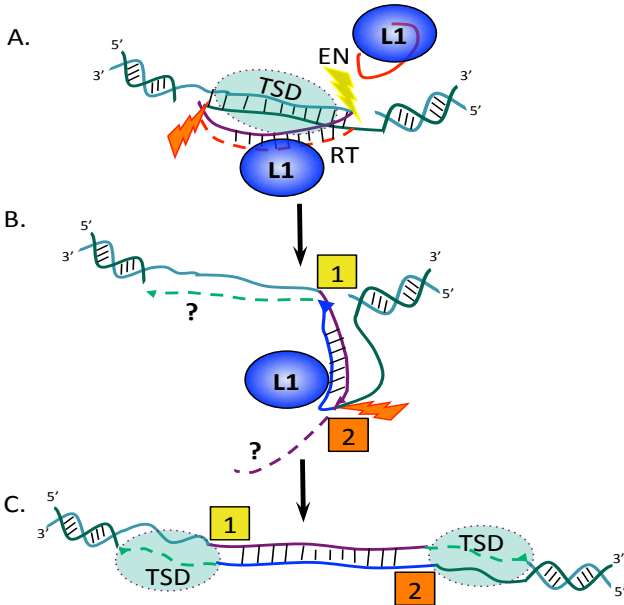


Figura 49. Mecanismo de retrotransposición de L1 durante la interfase celular. A) Tras el primer corte (rayo amarillo y 1), L1 polimeriza la primera cadena de DNA (en morado) a partir de su RNA (en rojo). **B)** Posteriormente se produce un segundo corte aguas arriba del primero (rayo naranja y 2), sobre el que L1 polimeriza la cadena complementaria (en azul) a su propio DNA (en morado). **C.** Como resultado, las dos cadenas de DNA permanecen unidas por el retroelemento, y a los lados cada una de las cadenas sencillas de la región genómica, que al copiarse forman las TSDs. La primera incógnita “?” indica que a la hora de cambiar de molde se pueda empezar a copiar aguas arriba de la primera cadena sintetizada y no al final (línea punteada morada), la segunda incógnita “?” es que es posible que al sintetizarse la cadena complementaria se complete también la primera TSD.

6.4.4.- Deleciones y duplicaciones largas en el sitio de inserción: mecanismos moleculares.

Normalmente el retroelemento produce dos cortes sencillos en posiciones desplazadas en cada hebra de DNA, dejando una región que posteriormente conformará las duplicaciones características (TSDs) de las inserciones del elemento (figura 49). Los huecos de cadena sencilla resultantes de la integración podrían ser rellenados por el propio L1 o por una polimerasa celular fácilmente, dando lugar a las TSDs y finalizando el evento de retrotransposición.

Nuestros datos revelan que existe una clara asociación entre las características anómalas y el defecto presente en las células FA, y apuntan a que dicha asociación parece producirse principalmente en eventos de retrotransposición que tienen lugar durante la replicación, en tanto que:

- La vía FA actúa principalmente durante la replicación del DNA.
- En las células FA, las inserciones anómalas se producen tanto en elementos mutantes en el dominio endonucleasa como en elementos salvajes.

Así, para explicar la aparición de las deleciones cortas encontradas al analizar nuestras inserciones, bastaría con que un elemento L1 accediera al lugar de la lesión y llevara a cabo su reversotranscripción, mientras que la maquinaria de la vía de escisión de nucleótidos eliminaría el oligonucleótido dañado, sobre el que la síntesis translesión no pudo actuar, y por tanto eliminando parte de la información genética original (figura 50).

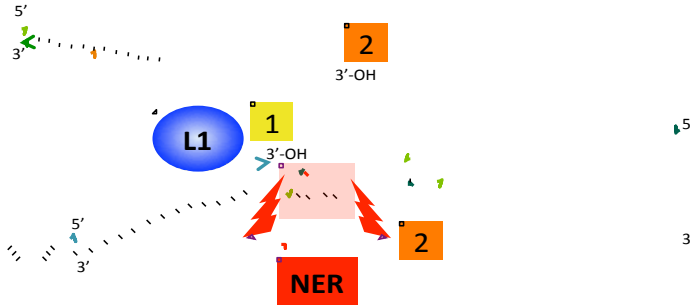


Figura 50. Modelo del mecanismo de aparición de deleciones cortas ante las inserciones de L1.

La incapacidad de reclutar a una polimerasa capaz de realizar la síntesis translesión puede permitir el acceso de L1 al extremo 3'-OH libre, mientras que la maquinaria de escisión de nucleótidos elimina el oligonucleótido dañado, de unas pocas pares de bases. No se requiere la participación de la endonucleasa de L1 para que se produzcan estos eventos, la posición de inicio de reversotranscripción podría empezar en la posición indicada como 1 en amarillo, mientras que la cadena complementaria se sintetizaría en una de las dos posiciones indicadas por el 2 naranja, que indican el otro extremo de la horquilla rota, o bien el extremo 3'-OH liberado tras la actuación de la vía de escisión de nucleótidos (NER), indicada en rojo. La NER eliminaría el pequeño fragmento de doble cadena completo, que al no haber sido replicado, se perdería de manera permanente.

Por otra parte, las duplicaciones inusualmente largas podrían estar produciéndose de una manera muy similar a lo que ocurre con las inserciones normales de L1, pero utilizando una posición más alejada que aparece debido al funcionamiento incorrecto de las proteínas afectadas en parte de la vía FA, o incluso aprovechando el funcionamiento de la maquinaria de la recombinación homóloga, una de las posibilidades por las que podrían aparecer estas duplicaciones se describe en la figura 51.

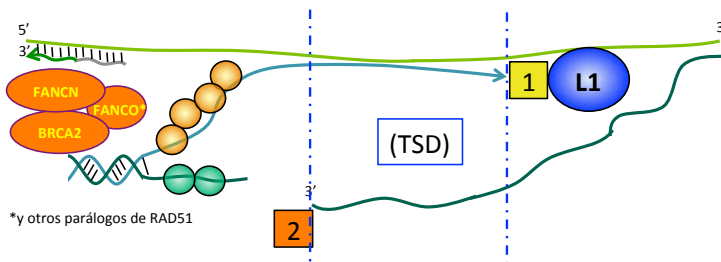


Figura 51. Modelo de mecanismo de aparición de duplicaciones largas del sitio de inserción ante las inserciones de L1.

La incapacidad de coordinar correctamente la maquinaria de la recombinación homóloga (o la SCE), y/o de reclutar una resolvasa, con el fin de promover y/o resolver las invasiones de cadena (o uniones de Hollyday), que se crean durante la recombinación homóloga normal, podría permitir el acceso de L1 al extremo 3'-OH libre de la cadena "invasora", indicado como 1 en amarillo, para polimerizar su DNA, y posteriormente alcanzar el otro extremo 3'-OH libre de la horquilla colapsada, indicado como 2 en naranja, para realizar la síntesis de la cadena complementaria. Esto permitiría la generación de TSDs de gran longitud, demarcadas entre las dos líneas azules verticales, incluso sin la participación de la endonucleasa de L1, de una manera análoga a como se generan las TSDs normalmente.

Aun así, debido a la gran cantidad de situaciones posibles ante una desregulación de la reparación asociada a la replicación (figura 52), la resolución de cualquier intento de integración por parte de L1 es igualmente variable ante deficiencias en la FA; por ejemplo, resulta muy difícil determinar la procedencia de las adiciones de nucleótidos sin molde aparente. Aunque existen varias posibilidades, puede que sea por la intervención de la maquinaria de reparación por unión de extremos no homólogos (Onozawa et al., 2014) o que sea la maquinaria de la recombinación homóloga, de una manera similar a

como se producen las duplicaciones segmentales al recombinarse dos secuencias similares de sitios distintos, quizá de una manera similar a lo que ocurre con las duplicaciones segmentales y Alu (Bailey et al., 2003). Esto último explicaría lo que ocurre en la inserción recuperada número 39, presentada previamente (figura 43, apartado 5.3.3), en la que dentro de la secuencia de relleno se aprecian dos regiones duplicadas que pudieron servir de homología para resolver la inserción.

Sin embargo, aunque no disponemos de datos suficientes para asegurar ni la causa ni la procedencia de estas largas secuencias de DNA de relleno sin molde, hemos podido encontrarlas ante deficiencias en FANCD2, FANCI y FANCC, lo que sugiere que su aparición está también directamente relacionada con el fallo de la ruta FA.

Por otra parte, dado que la actividad de los elementos salvajes no parece verse impedida o gravemente disminuida en ninguna de las líneas deficientes en la vía FA estudiadas o siRNA empleados (a excepción de PCNA), al igual que ocurre en las líneas deficientes en la vía NHEJ, cabe pensar que el proceso de retrotransposición mediado por TPRT no requiere para completarse de la participación de la recombinación homóloga, la vía de unión de extremos homólogos o la síntesis translesión; sin embargo, la drástica bajada de retrotransposición al inhibir PCNA podría indicar que la maquinaria de replicación sí es necesaria para completar los eventos de retrotransposición, como recientemente se ha propuesto en un estudio sobre las interacciones de L1 con la maquinaria celular (Taylor et al., 2013).

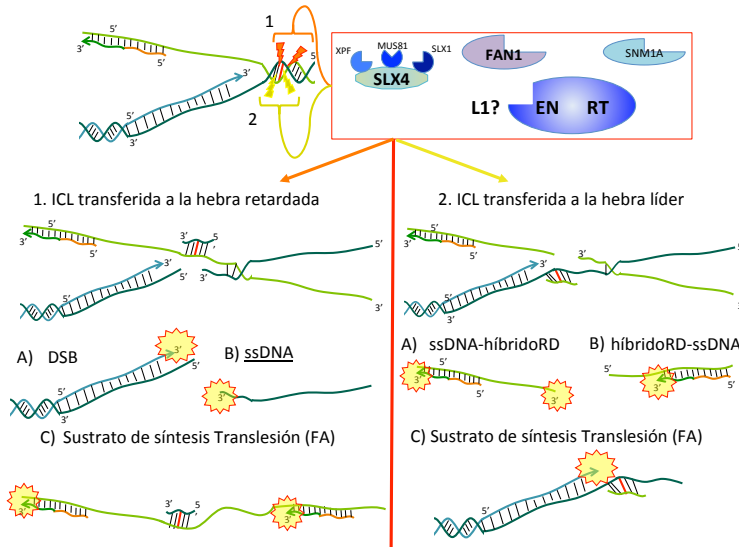


Figura 52. Estructuras intermedias de la reparación de horquillas de replicación colapsadas.

Durante la reparación de daños como los ICLs, la primera parte de la reparación implica diferentes cortes en el DNA que determinaran a qué hebra se desplaza el daño y qué hebra se utilizara para reparar la lesión. Estos cortes, mediados por las nucleasas FAN1 o las asociadas con FANCP (SLX4), liberan extremos 3'-OH (indicados mediante una estrella amarilla), que podrían ser utilizados por los elementos L1 deficientes en su endonucleasa. Cada uno de estos extremos hidroxilo libres en 3' supone una posibilidad para la vía de retrotransposición ENi en las células deficientes en la vía FA, con diferentes resultados posibles. También existe la posibilidad de que, en el caso de retroelementos salvajes, la propia endonucleasa de L1 pueda realizar uno de los dos cortes iniciales, indicados como rayos amarillos (cortes en la hebra retardada) o naranjas (cortes en la hebra líder).

6.5.- Posibles implicaciones de la actividad de L1 en la progresión de la FA y la oncogénesis.

Como se ha mencionado previamente, la característica más común en los pacientes de FA es la elevada incidencia de leucemias y tumores sólidos, que surgen de la acumulación de mutaciones en las células de dichos pacientes (revisado en Alter, 2003; Moldovan & D'Andrea, 2009). La elevada inestabilidad genómica, así como la alta predisposición a tumores, se ha considerado siempre como consecuencia de la reparación anómala que se produce en estas células ante las diferentes fuentes de daño, pero las fuentes de daño consideradas son principalmente de origen exógeno, tales como la exposición a agentes intercalantes y estrés oxidativo (revisado en Kee & D'Andrea, 2012).

En este estudio, por primera vez, hemos analizado la respuesta de las células de estos pacientes ante el mayor agente mutagénico intrínseco al genoma humano, el retroelemento L1, cuya actividad ya se ha demostrado como causa directa de tumores y cáncer (Iskow et al, 2010; Lee et al., Science 2012; Soylo et al., Genome Res 2012; Shukla et al, 2013; Helman et al., 2014; revisado en Carreira et al., 2013). Además, hemos puesto de manifiesto que existe una importante desregulación tanto del nivel de actividad como de las características de las inserciones de LINE-1 en las células de pacientes de anemia de Fanconi.

Además, las deficiencias de algunos miembros, son especialmente severas y provocan la aparición de

diversos tumores, destacando el cáncer de células escamosas de cabeza y cuello (*Head and Neck Squamous Cell Carcinoma* o HNSCC) o el cáncer de mama (o BRCA). Además, curiosamente, un estudio reciente ha encontrado una alta incidencia de inserciones somáticas de L1 en estos tipos de cáncer (Helman et al., 2014); paralelamente, en nuestro estudio, la inhibición mediante siRNAs de genes FA que también están implicados en la recombinación homóloga, concretamente BRCA2, FANCI, FANCD1 y FANCD2, apareció asociada a un incremento notable de la actividad de L1, con la excepción de FANCD3, que presentó una modesta reducción de la movilidad, si bien, hemos de destacar dos factores importantes (revisado en Kee & D'Andrea, 2012):

- Los portadores heterocigotos de mutaciones en BRCA2, FANCI, FANCD1 y FANCD2 presentan una predisposición aumentada a la aparición de cáncer de mama y ovario (Rosemberg et al., 2003).

- Se ha descrito que la participación de FANCD3 en la recombinación homóloga parece más asociada a mediar la comunicación de HR y FA, y, si bien los pacientes de FA con mutaciones en FANCD3 presentan un fenotipo severo, no se ha descrito asociación de portadores heterocigotos de mutaciones en FANCD3 con predisposición a cáncer de mama (Tebbs et al., 2005; Hussain et al., 2006; Wilson et al., 2008).

Asimismo, otro estudio reciente ha demostrado la participación de L1 en la reparación de roturas de doble cadena, encontrando varios ejemplos en el genoma humano (Onozawa et al., 2014). Las deficiencias en los mecanismos de reparación de las roturas de doble cadena son una

conocida causa de inestabilidad genética y cáncer (revisado en Deriano & Roth, 2013; Burrell et al., 2013), y nuestros datos sugieren que se produce un incremento dramático de la actividad de L1 ante fallos en la reparación de estas roturas.

Si tenemos en cuenta observaciones previas, nuestro estudio pone de manifiesto que L1 es sin duda una de las posibles causas de malignización celular en la anemia de Fanconi y con una probable implicación de su actividad en la progresión de los tumores. La actividad de L1 podría estar jugando un papel muy importante en el desarrollo de líneas celulares tumorales subclonales resistentes a los tratamientos actualmente utilizados en la clínica. El motivo es que estos tratamientos siguen principalmente dos estrategias: inducir daños en el DNA, mediante radioterapia o quimioterapia o, inhibir la maquinaria de reparación celular (revisado en Curtin, 2012).

6.6.- Posibilidad de tratamiento expandido de la FA.

Como se ha mencionado anteriormente, la complejidad en la clínica de la FA se pone realmente de manifiesto a la hora de seleccionar un tratamiento adecuado para los tumores o las leucemias que acaban desarrollando estos pacientes (revisado en Kee & D'Andrea, 2009). Esto se debe a que, como se decía anteriormente, la mayoría de los tratamientos antitumorales están basados en producir daños en el DNA, para así interferir con la replicación y de alguna manera controlar la proliferación desmesurada de las células cancerígenas (revisado en Curtin, 2012). Sin

embargo, en el caso de la anemia de Fanconi ocurre lo contrario: las células características de las anemias aplásicas de estos pacientes presentan una carencia de funcionalidad, así como dificultades para la proliferación; todo esto, unido a la extrema sensibilidad a los agentes intercalantes del DNA, tales como la Mitomicina C o el cisplatino, complica enormemente la elección de un tratamiento adecuado, ya que como mencionábamos anteriormente, muchos de los agentes quimioterápicos que se utilizan en la clínica son concretamente agentes intercalantes del DNA (revisado en Curtin, 2012).

Con vistas al futuro, se estudian varias posibilidades de tratamiento, principalmente la obtención de células del paciente con el fin de generar células madre hematopoyéticas (*Hematopoietic Stem Sells* o HSCs) sanas (revisado en Kee & D'Andrea, 2012).

Se han generado células madre reprogramadas pluripotentes (*induced Pluripotent Stem Cells* o iPSCs), con objeto de corregir mediante terapia génica el gen deficiente, diferenciarlas a HSCs y reintroducirlas en el paciente (Raya et al., 2009). El problema de este tratamiento es que las estrategias de terapia génica diseñadas hace años estaban basadas en vectores virales que poseían cierta preferencia por regiones transcripcionalmente activas y otras características que podrían promover la transformación oncogénica de la célula (Shou et al., 2006). Actualmente las estrategias de terapia génica y celular han mejorado mucho, tanto en seguridad para el paciente como en eficiencia para la corrección del defecto, así encontramos:

-Vectores cada vez más seguros, tanto para reprogramación como para terapia génica (Grabundzija et al., 2013; Molina-Estevez et al., 2013).

-Mejores protocolos de obtención, mantenimiento y manipulación de las líneas celulares de la FA (Müller et al., 2012; Amabile et al., 2013).

-La capacidad de comprobar donde se produce el evento de integración del gen introducido y la integridad del genoma celular antes de reintroducir las células corregidas en el paciente (Howden et al., 2011).

Sin embargo, a pesar de todas las ventajas que supondría este tratamiento, el problema es que, al igual que el trasplante de médula ósea (BMT) o incluso la propia reversión fenotípica de la enfermedad, que da como resultado mosaicismo, no previene la aparición de tumores (Gregory et al., 2001). Actualmente resultaría técnicamente muy difícil alcanzar otros tejidos para repoblarlos con el fin de evitar esta posibilidad.

Es aquí donde puede residir la importancia de los datos recogidos en esta tesis, ya que nuestros experimentos indican que una disminución de la actividad de L1 en estos pacientes podría reducir la carga de mutaciones que padecen, y por tanto las posibilidades de desarrollar tumores. Para ello, existen compuestos que presentan actividad frente a las reversotranscriptasas, que se diseñaron originalmente frente a la RT del virus HIV y que, de hecho, se utilizan a día de hoy en la clínica para combatir el SIDA, que presentan además actividad frente a la RT del L1 (Jones et al., 2008; Dai et al, 2011 y datos de laboratorio no publicados). A la luz de nuestros resultados creemos que el uso de estos

inhibidores de RT en los pacientes de anemia de Fanconi, y posiblemente otros síndromes de inestabilidad genética, podría disminuir el riesgo de desarrollar tumores. Obviamente, el primer paso sería asegurar su seguridad y eficacia en células indiferenciadas y modelos murinos que presenten el defecto, con la ventaja de que ya se conoce que estos compuestos inhibidores de la actividad RT no son tóxicos para el ser humano.

CONCLUSIONES

7.- CONCLUSIONES.

1- Las deficiencias en algunos miembros de la vía de reparación de la Anemia de Fanconi aumentan la frecuencia de retrotransposición de L1 por la vía independiente de endonucleasa.

2- En los eventos de retrotransposición de los elementos L1 estudiados en células deficientes en la vía de reparación de la anemia de Fanconi, las inserciones presentan características anómalas como duplicaciones excesivamente largas o deleciones. Estas características se asocian a procesos de inestabilidad genómica.

3- Los niveles de retrotransposición del elemento L1 varían al disminuir los niveles de expresión de las proteínas de la vía de reparación de la anemia de Fanconi.

4- Los niveles de retrotransposición del elemento L1 disminuyen al disminuir los niveles de la proteína PCNA.

5- La maquinaria de reparación y replicación participa durante el proceso de retrotransposición de L1 y/o en la resolución de sus intentos de integración.

BIBLIOGRAFÍA

8.- BIBLIOGRAFÍA.

- Adamo A, Collis SJ, Adelman CA, Silva N, Horejsi Z, Ward JD, Martinez-Perez E, Boulton SJ, La Volpe A. 2010. Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. . Mol Cell. 39(1):25-35.
- Adams JW, Kaufman RE, Kretschmer PJ, Harrison M, Nienhuis AW. 1980. A family of long reiterated DNA sequences, one copy of which is next to the human beta globin gene. Nucleic Acids Res. 8(24):6113-28.
- Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV. 2006. Unconventional translation of mammalian LINE-1 retrotransposons. Genes Dev. 20:210-24.
- Alter BP. 2003. Cancer in Fanconi anemia, 1927-2001. Cancer. 97(2):425-40.
- Alter BP, Rosenberg PS, Brody LC. 2007. Clinical and molecular features associated with biallelic mutations in FANCD1/BRCA2. J Med Genet. 44(1):1-9.
- Alter BP, Giri N, Savage SA, Peters JA, Loud JT, Leathwood L, Carr AG, Greene MH, Rosenberg PS. 2010. Malignancies and survival patterns in the National Cancer Institute inherited bone marrow failure syndromes cohort study. Br J Haematol. 150(2):179-88.
- Amabile G, Welner RS, Nombela-Arrieta C, D'Alise AM, Di Ruscio A, Ebralidze AK, Kravtsov Y, Ye M, Kocher O, Neuberg DS, Khrapko K, Silberstein LE, Tenen DG. 2013. In vivo generation of transplantable human hematopoietic cells from induced pluripotent stem cells. Blood. 121(8):1255-64.

- Ameziane N, Errami A, Léveillé F, Fontaine C, de Vries Y, van Spaendonk RM, de Winter JP, Pals G, Joenje H. 2008. Genetic subtyping of Fanconi anemia by comprehensive mutation screening. *Hum Mutat.* 29:159–66.
- Andreassen PR, D'Andrea AD, Taniguchi T. 2004. ATR couples FANCD2 monoubiquitination to the DNA-damage response. *Genes Dev.* 18(16):1958-63.
- Aplan, PD. 2006. Chromosomal translocations involving the MLL gene: molecular mechanisms. *DNA Repair (Amst).* 5(9-10):1265-72.
- Apostolou S, Whitmore SA, Crawford J, Lennon G, Sutherland GR, Callen DF, Lanzano L, Savino M, D'Apolito M, Notarangelo A, Memeo E et al. 1996. Fanconi anaemia/Breast cancer consortium. Positional cloning of the Fanconi anaemia group A gene. *Nat Genet.* 14:324–8.
- Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ. 2007. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science.* 316(5825):744-7.
- Arjan-Odedra S, Swanson CM, Sherer NM, Wolinsky SM, Malim MH. 2012. Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. *Retrovirology* 9:53.
- Athanikar JN, Badge RM, Moran JV. 2004. A YY1-binding site is required for accurate human LINE-1 transcription initiation. *Nucleic Acids Res.* 32:3846–55.
- Auerbach AD. 1993. Fanconi anemia diagnosis and the diepoxybutane (DEB) test. *Exp Hematol.* 21(6):731–733.

- Avery OT, Macleod CM, McCarty M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types : induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J Exp Med.* 79(2):137-58.
- Babushok DV, Kazazian HH Jr. 2007. Progress in understanding the biology of the human mutagen LINE-1. *Hum Mutat.* 28(6):527-39.
- Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA et al. 2011. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 479:534–37.
- Bailey JA, Eichler EE. 2006. Primate segmental duplications: crucibles of evolution, diversity and disease. *Nat Rev Genet.* 7(7):552-64.
- Bancroft I, Dean C. 1993. Factors affecting the excision frequency of the maize transposable element Ds in *Arabidopsis thaliana*. *Mol Gen Genet.* 240(1):65-72.
- Bogliolo M, Cabré O, Callén E, Castillo V, Creus A, Marcos R, Surrallés J. 2002. The Fanconi anaemia genome stability and tumour suppressor network. *Mutagenesis* 17(6):529-38.
- Beck CR, Collier P, Macfarlane C, Malig M, Kidd JM et al. 2010. LINE-1 retrotransposition activity in human genomes. *Cell* 141:1159–70.
- Beck CR, Garcia-Perez JL, Badge RM, Moran JV. 2011. LINE-1 elements in structural variation and disease. *Annu. Rev. Genomics Hum. Genet.* 12:187–215.
- Becker KG, Swergold GD, Ozato K, Thayer RE. 1993. Binding of the ubiquitous nuclear transcription factor YY1 to a cis regulatory sequence in the human LINE-1 transposable element. *Hum. Mol. Genet.* 2:1697–702.

- Belancio VP, Roy-Engel AM, Pochampally RR, Deininger P. 2010. Somatic expression of LINE-1 elements in human tissues. *Nucleic Acids Res.* 38(12):3909-22
- Belfort M, Derbyshire V, Parker MM, Cousineau B, Lambowitz AM. 2002. Mobile introns: pathways and proteins. In: Craig N, Craigie R, Gellert M, Lambowitz AM, editors. *Mobile DNA II*. Washington DC: ASM Press. pp. 761–783.
- Blom E, van de Vrugt HJ, de Vries Y, de Winter JP, Arwert F, Joenje H. 2004. Multiple TPR motifs characterize the Fanconi anemia FANCG protein. *DNA Repair (Amst)*. 3(1):77-84.
- Blot M. 1994. *Genetica*. Transposable elements and adaptation of host bacteria. 93(1-3):5-12.
- Bogerd HP, Wiegand HL, Hulme AE, Garcia-Perez JL, O'Shea KS et al. 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc. Natl. Acad. Sci. USA* 103:8780–85.
- Boyd EF, Hartl DL. 1997. Nonrandom location of IS1 elements in the genomes of natural isolates of *Escherichia coli*. *Mol Biol Evol.* 14(7):725-32.
- Branzei D, Foiani M. 2010. Maintaining genome stability at the replication fork. *Nat Rev Mol Cell Biol.* 11(3):208-19.
- Britten RJ, Kohne DE. 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science.* 161(3841):529-40.
- Britten RJ. 1996. DNA sequence insertion and evolutionary variation in gene regulation. *Proc Natl Acad Sci U S A.* 93(18):9374-7.
- Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH et al. 2003. Hot L1s account for the bulk of

retrotransposition in the human population. *Proc. Natl. Acad. Sci. USA* 100:5280–85.

- Burrell RA, McGranahan N, Bartek J, Swanton C. 2013. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*. 501(7467):338-45.

- Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S et al. 2001. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell*. 105(1):149-60.

- Carreau M, Alon N, Bosnoyan-Collins L, Joenje H, Buchwald M. 1999. Drug sensitivity spectra in Fanconi anemia lymphoblastoid cell lines of defined complementation groups. *Mutat Res*. 435: 103–109.

- Carreira PE, Richardson SR, Faulkner GJ. 2014. L1 retrotransposons, cancer stem cells and oncogenesis. *FEBS J*. 281(1):63-73.

- Castella M, Pujol R, Callén E, Trujillo JP, Casado JA, Gille H, Lach FP, Auerbach AD et al, 2011. Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. *Blood*. 117(14):3759-69.

- Chandra S, Levran O, Jurickova I, Maas C, Kapur R, Schindler D, Henry R, Milton K, Batish SD, Cancelas JA, Hanenberg H, Auerbach AD, Williams DA. 2005. A rapid method for retrovirus-mediated identification of complementation groups in Fanconi anemia patients. *Mol Ther*. 12:976–84.

- Christensen SM, Eickbush TH. 2005. R2 target-primed reverse transcription: ordered cleavage and polymerization steps by protein subunits asymmetrically bound to the target DNA. *Mol Cell Biol*. 25(15):6617-28.

- Clements AP, Singer MF. 1998. The human LINE-1 reverse transcriptase:effect of deletions outside the common reverse transcriptase domain. *Nucleic Acids Res.* 1998 26(15):3528-35.
- Collins NB, Wilson JB, Bush T, Thomashevski A, Roberts KJ, Jones NJ, Kupfer GM. 2009. ATR-dependent phosphorylation of FANCA on serine 1449 after DNA damage is important for FA pathway function. *Blood.* 113(10):2181-90.
- Cordaux R. 2009. Gene conversion maintains nonfunctional transposable elements in an obligate mutualistic endosymbiont. *Mol Biol Evol.* 26(8):1679-82.
- Cordaux R, Batzer MA. 2009. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* 10:691-703.
- Cost GJ, Boeke JD. 1998. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* 37:18081-93.
- Cost GJ, Golding A, Schlissel MS, Boeke JD. 2001. Target DNA chromatinization modulates nicking by L1 endonuclease. *Nucleic Acids Res.* 29:573-77.
- Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, et al. 2009. L1 retrotransposition in human neural progenitor cells. *Nature* 460:1127-31.
- Coufal NG, Garcia-Perez JL, Peng GE, Marchetto MC, Muotri AR, et al. 2011. Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc. Natl. Acad. Sci. USA* 108:20382-87
- Craig NL, Craigie R, Gellert M, Lambowitz AM. 2002. *Mobile DNA II.* Washington, DC: Am. Soc. Microbiol (ASM Press).

- Curtin NJ. 2012. DNA repair dysregulation from cancer driver to therapeutic target. *Nat Rev Cancer*. 12(12):801-17.
- D'Andrea AD. 2010. Susceptibility pathways in Fanconi's anemia and breast cancer. *N Engl J Med*. 362(20):1909-1919.
- Dai L1, Huang Q, Boeke JD. 2011. Effect of reverse transcriptase inhibitors on LINE-1 and Ty1 reverse transcriptase activities and on LINE-1 retrotransposition. *BMC Biochem*. 12:18.
- Damert A, Raiz J, Horn AV, Löwer J, Wang H, Xing J et al. 2009. 5'-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. *Genome Res*. 19(11):1992-2008.
- Daniels MJ, Wang Y, Lee M, Venkitaraman AR. 2004. Abnormal cytokinesis in cells deficient in the breast cancer susceptibility protein BRCA2. *Science*. 306(5697):876-9.
- Darwin C. 1859. *On the Origin of Species*. London: Murray
- Davies AA, Masson JY, McIlwraith MJ, Stasiak AZ et al. 2001. Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol Cell*. 7(2):273-82.
- de Boer M, van Leeuwen K, Geissler J, Weemaes CM, van den Berg TK, et al. 2014. Primary immunodeficiency caused by an exonized retroposed gene copy inserted in the CYBB gene. *Hum. Mutat*. 35:486-96.
- de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. 2011. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet*. 7:e1002384.

- Deans AJ, West SC. 2011. DNA interstrand crosslink repair and cancer. *Nat Rev Cancer*. 11(7):467-80.
- Delattre M, Tatout C, Coen D. 2000. P-element transposition in *Drosophila melanogaster*: influence of size and arrangement in pairs. *Mol Gen Genet*. 263(3):445-54.
- Dewannieux M, Esnault C, Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet*. 35(1):41-8.
- Dias ES, Carareto CM. 2012. Ancestral polymorphism and recent invasion of transposable elements in *Drosophila* species. *BMC Evol Biol*. 23;12:119.
- DiGiovanni L, Haynes SR, Misra R, Jelinek WR. 1983. Kpn I family of long-dispersed repeated DNA sequences of man: evidence for entry into genomic DNA of DNA copies of poly(A)-terminated Kpn I RNAs. *Proc Natl Acad Sci U S A*. 80(21):6533-7.
- Dmitriev SE, Andreev DE, Terenin IM, Olovnikov IA, Prassolov VS, et al. 2007. Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5' untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal ribosome entry site mediated. *Mol. Cell. Biol*. 27:4685-97.
- Dombroski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HH Jr. 1991. Isolation of an active human transposable element. *Science* 254:1805-8.
- Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature*. 284(5757):601-3.
- Doucet AJ, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, et al. 2010. Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet*. 6:e1001150.

- Dunham I. 2002. Human genome sequences: enigmatic variations. *Mutagenesis*. 17(6):457-61.
- Engels WR, Johnson-Schlitz DM, Eggleston WB, Sved J. 1990. High-frequency P element loss in *Drosophila* is homolog dependent. *Cell*. 62(3):515-25.
- Esnault C, Maestre J, Heidmann T. 2000. Human LINE retrotransposons generate processed pseudogenes. *Nat. Genet*. 24:363-67.
- Ewing AD, Kazazian HH Jr. 2010. High-throughput sequencing reveals extensive variation in humanspecific L1 content in individual human genomes. *Genome Res*. 20:1262-70.
- Fagerlie SR, Koretsky T, Torok-Storb B, Bagby GC. 2004. Impaired type I IFN-induced Jak/STAT signaling in FA-C cells and abnormal CD4+ Th cell subsets in *Fancc*^{-/-} mice. *J Immunol*. 173(6):3863-70.
- Ferrer M, Rodríguez JA, Spierings EA, de Winter JP, Giaccone G, Kruyt FA. 2005. Identification of multiple nuclear export sequences in Fanconi anemia group A protein that contribute to CRM1-dependent nuclear export. *Hum Mol Genet*. 14(10):1271-81
- Feschotte C, Pritham EJ. 2007. DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*. 41:331-68.
- Fiers W, Contreras R, Duerinck F, Haegeman G, Iserentant D et al. 1976. Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature*. 260(5551):500-7.
- Frost RJ, Hamra FK, Richardson JA, Qi X, Bassel-Duby R, Olson EN. 2010. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc Natl Acad Sci U S A*. 107(26):11847-52.

- Fu D, Dudimah FD, Zhang J, Pickering A, Paneerselvam J et al. 2013. Recruitment of DNA polymerase eta by FANCD2 in the early response to DNA damage. *Cell Cycle*. 12(5):803-9.
- Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H et al. 2013. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat Struct Mol Biol*. 20(3):332-8.
- Fanconi, G. 1927. Familiaere infantile perniziosaartige Anaemie (pernioeses Blutbild and konstitution). *Jahrbuch Kinderheild* 117, 257–280.
- Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, et al. 2009. The regulated retrotransposon transcriptome of mammalian cells. *Nat. Genet*. 41:563–71.
- Feng Q, Moran JV, Kazazian HH Jr, Boeke JD. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87:905–16.
- Freeman JD, Goodchild NL, Mager DL. 1994. A modified indicator gene for selection of retrotransposition events in mammalian cells. *Biotechniques*. 17:47-52.
- Fujita A, Hattori M, Takenaka O, Sakaki Y. 1987. The L1 family (KpnI family) sequence near the 3' end of human beta-globin gene may have been derived from an active L1 sequence. *Nucleic Acids Res*. 15(10):4007-20.
- Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, Grompe M, D'Andrea AD. 2001. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell*. 7:249.

- Garcia-Perez JL, Marchetto MC, Muotri AR, Coufal NG, Gage FH, et al. 2007. LINE-1 retrotransposition in human embryonic stem cells. *Hum. Mol. Genet.* 16:1569–77.
- Garcia-Perez JL, Morell M, Scheys JO, Kulpa DA, Morell S, et al. 2010. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. *Nature* 466:769–73.
- Gardner RS, Wahba AJ, Basilio C, Miller RS, Lengyel P, Speyer JF. 1962. Synthetic polynucleotides and the amino acid code. VII. *Proc Natl Acad Sci U S A.* 48:2087-94.
- Gilbert N, Lutz-Prigge S, Moran JV. 2002. Genomic deletions created upon LINE-1 retrotransposition. *Cell* 110:315–25.
- Gilbert N, Lutz S, Morrish TA, Moran JV. 2005. Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol. Cell. Biol.* 25:7780–95.
- Goodier JL, Ostertag EM, Kazazian HH Jr. 2000. Transduction of 3'-flanking sequences is common in L1 retrotransposition. *Hum. Mol. Genet.* 9:653–57.
- Goodier JL, Kazazian HH Jr. 2008. Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* 135:23–35
- Goodier JL, Cheung LE, Kazazian HH Jr. 2012. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genet.* 8:e1002941.
- Goodier JL, Cheung LE, Kazazian HH Jr. 2013. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res.* 41:7401–19.
- Grabundzija I, Wang J, Sebe A, Erdei Z, Kajdi R, Devaraj A, Steinemann D, Szuhai K, Stein U, Cantz T, Schambach A, Baum C, Izsvák Z, Sarkadi B, Ivics Z.

2013. Sleeping Beauty transposon-based system for cellular reprogramming and targeted gene insertion in induced pluripotent stem cells. *Nucleic Acids Res.* 41(3):1829-47.

- Grandi FC, An W. 2013. Non-LTR retrotransposons and microsatellites: Partners in genomic variation. *Mob Genet Elements.* 3(4):e25674.

- Gregory JJ Jr, Wagner JE, Verlander PC, Levran O, Batish SD, Eide CR, Steffenhagen A, Hirsch B, Auerbach AD. 2001. Somatic mosaicism in Fanconi anemia: evidence of genotypic reversion in lymphohematopoietic stem cells. *Proc Natl Acad Sci USA.* 98(5):2532-2537.

- Griffith F. 1928. The Significance of Pneumococcal Types. *J Hyg (Lond).* 27(2):113-59.

- Grimaldi G, Singer MF. 1983. Members of the KpnI family of long interspersed repeated sequences join and interrupt alpha-satellite in the monkey genome. *Nucleic Acids Res.* 11(2):321-38.

- Grompe M, D'Andrea A. 2001. Fanconi anemia and DNA repair. *Hum Mol Genet.* 10(20):2253-9.

- Gross M, Hanenberg H, Lobitz S, Friedl R, Herterich S, Dietrich R et al. 2002. Reverse mosaicism in Fanconi anemia: natural gene therapy via molecular self-correction. *Cytogenet Genome Res.* 98(2-3):126-35.

- Han JS, Szak ST, Boeke JD. 2004. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* 429:268-74.

- Hancks DC, Kazazian HH Jr. 2012. Active human retrotransposons: variation and disease. *Curr. Opin. Genet. Dev.* 22:191-203.

- Hartl DL, Lohe AR, Lozovskaya ER. 1997. Modern thoughts on an ancient mariner: function, evolution, regulation. *Annu Rev Genet* 31:337–58.
- Hata K, Sakaki Y. 1997. Identification of critical CpG sites for repression of L1 transcription by DNA methylation. *Gene*. 189(2):227-34.
- Hatzopoulos P1, Monastirioti M, Yannopoulos G, Louis C. 1987. The instability of the TE-like mutation Dp(2:2)GYL of *Drosophila melanogaster* is intimately associated with the hobo element. *EMBO J*. 6(10):3091-96.
- Havecker ER, Gao X, Voytas DF. 2004. The diversity of LTR retrotransposons. *Genome Biol*. 5(6):225.
- Heidmann, T., Heidmann, O., and Nicholas, J.F. 1988. An indicator gene to demonstrate intracellular transposition of defective retroviruses. *Proc. Natl. Acad. Sci. USA* 85, 2219–2223.
- Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Res*. 24(7):1053-63.
- Heras SR, Macias S, Plass M, Fernandez N, Cano D, Eyraas E, Garcia-Perez JL, Cáceres JF. 2013. The Microprocessor controls the activity of mammalian retrotransposons. *Nat Struct Mol Biol*. 20(10):1173-1181.
- Hohjoh H, Singer MF. 1996. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *EMBO J*. 15:630–39.
- Holmes SE, Singer MF, Swergold GD. Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element. *J Biol Chem*. 267(28):19765-8.

- Horn AV, Klawitter S, Held U, Berger A, Vasudevan AA, Bock A et al. 2014. Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity. *Nucleic Acids Res.* 42(1):396-416.
- Howden SE, Gore A, Li Z, Fung HL, Nisler BS, Nie J, Chen G, McIntosh BE, Gulbranson DR, Diol NR, Taapken SM, Vereide DT, Montgomery KD, Zhang K, Gamm DM, Thomson JA. 2011. Genetic correction and analysis of induced pluripotent stem cells from a patient with gyrate atrophy. *Proc Natl Acad Sci USA.* 108(16):6537-42.
- Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q et al. 2002. Biallelic inactivation of BRCA2 in Fanconi anemia. *Science.* 297(5581):606-9.
- Howlett NG, Harney JA, Rego MA, Kolling FW 4th, Glover TW. 2009. Functional interaction between the Fanconi Anemia D2 protein and proliferating cell nuclear antigen (PCNA) via a conserved putative PCNA interaction motif. *J Biol Chem.* 284(42):28935-42.
- Hu W, Das OP, Messing J. 1995. Zeon-1, a member of a new maize retrotransposon family. *Mol Gen Genet.* 248:471-480.
- Hu XY, Ray PN, Worton RG. 1991. Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO J.* 10(9):2471-7.
- Huang M, D'Andrea AD. 2010. A new nuclease member of the FAN club. *Nat Struct Mol Biol.* 17(8):926-8.

- Huang M, Kim JM, Shiotani B, Yang K, Zou L, D'Andrea AD. 2010. The FANCM/FAAP24 complex is required for the DNA interstrand crosslink-induced checkpoint response. *Mol Cell*. 39(2):259-68.
- Hussain S, Wilson JB, Blom E, Thompson LH, Sung P et al. 2006. Tetratricopeptide-motif-mediated interaction of FANCG with recombination proteins XRCC3 and BRCA2. *DNA Repair (Amst)*. 5(5):629-40.
- Iskow RC, McCabe MT, Mills RE, Torene S, Pittard WS, et al. 2010. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* 141:1 253-61.
- Jiang N, Bao Z, Temnykh S, Cheng Z, Jiang J, Wing RA, McCouch SR, Wessler SR. 2002. Dasheng: a recently amplified nonautonomous long terminal repeat element that is a major component of pericentromeric regions in rice. *Genetics*. 161:1293-1305.
- Jiang N, Jordan IK, Wessler SR. 2002. Dasheng and RIRE2. A nonautonomous long terminal repeat element and its putative autonomous partner in the rice genome. *Plant Physiol*. 130:1697-1705.
- Jinks-Robertson S, Petes TD. 1986. Chromosomal translocations generated by high-frequency meiotic recombination between repeated yeast genes. *Genetics*. 114(3):731-52.
- Joenje, H, Patel, KJ. 2001. The emerging genetic and molecular basis of Fanconi anaemia. *Nature Rev. Genet*. 2: 446-457.
- Jurka J. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc. Natl. Acad. Sci. USA* 94:1872-77.
- Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, et al. 2009. L1 retrotransposition occurs mainly in

embryogenesis and creates somatic mosaicism. *Genes Dev.* 23:1303–12.

- Kapitonov VV, Jurka J. 2007. Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet.* 23(10):521-9.

- Kazazian HH Jr, Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* 332:164–66.

- Kazazian HH Jr, Scott AF. 1993. "Copy and paste" transposable elements in the human genome. *J Clin Invest.* 91(5):1859-60.

- Kazazian HH. 1998. Mobile elements and disease. *Jr Curr. Opin. Genet. Dev.* 8, 343–350.

- Kazazian HH Jr. 1999. An estimated frequency of endogenous insertional mutations in humans. *Nat Genet.* 22(2):130.

- Kazazian HH Jr, Goodier JL. 2002. LINE drive, retrotransposition and genome instability. *Cell.* 110(3):277-80.

- Kazazian HH Jr. 2004. Mobile elements: drivers of genome evolution. *Science.* 303(5664):1626-32.

- Kee Y, D'Andrea AD. 2012. Molecular pathogenesis and clinical management of Fanconi anemia. *J Clin Invest.* 122(11):3799-806.

- Kennedy N, Knedlitschek G, Groner B, Hynes NE, Herrlich P, Michalides R, van Ooyen AJ. 1982. Long terminal repeats of endogenous mouse mammary tumour virus contain a long open reading frame which extends into adjacent sequences. *Nature.* 295(5850):622-4.

- Kent WJ. 2002. BLAT--the BLAST-like alignment tool. *Genome Res.* 12(4):656-64.

- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res.* 12(6):996-1006.
- Khazina E, Weichenrieder O. 2009. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc. Natl. Acad. Sci. USA* 106:731–36.
- Khazina E, Truffault V, Büttner R, Schmidt S, Coles M, Weichenrieder O. 2011. Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. *Nat Struct Mol Biol.* 18(9):1006-14.
- Kim A, Terzian C, Santamaria P, Pelisson A, Prud'homme N, Bucheton A. 1994. Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *Proc Natl Acad Sci USA.* 91:1285-1289.
- Kim H, D'Andrea AD. 2012. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev.* 26(13):1393-408.
- Kim H, Yang K, Dejsuphong D, D'Andrea AD. 2012. Regulation of Rev1 by the Fanconi anemia core complex. *Nat Struct Mol Biol.* 19(2):164-70.
- Kim PM, Allen C, Wagener BM, Shen Z, Nickoloff JA. Overexpression of human RAD51 and RAD52 reduces double-strand break-induced homologous recombination in mammalian cells. *Nucleic Acids Res.* 29(21):4352-60.
- Kinomoto M, Kanno T, Shimura M, Ishizaka Y, Kojima A, et al. 2007. All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. *Nucleic Acids Res.* 35:2955–64.
- Knipscheer P, Räschle M, Smogorzewska A, Enoiu M, Ho TV et al. 2009. The Fanconi anemia pathway

promotes replication-dependent DNA interstrand cross-link repair. *Science*. 326(5960):1698-701.

- Kottemann MC, Smogorzewska A. 2013. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature*. 493(7432):356-63.

- Kubo S, Seleme MC, Soifer HS, Perez JL, Moran JV, et al. 2006. L1 retrotransposition in nondividing and primary human somatic cells. *Proc. Natl. Acad. Sci. USA* 103:8036-41.

- Kulpa DA, Moran JV. 2005. Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. *Hum. Mol. Genet.* 14:3237-48.

- Kulpa DA, Moran JV. 2006. cis-Preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat. Struct. Mol. Biol.* 13:655-60.

- Kutler DI, Auerbach AD. 2004. Fanconi anemia in Ashkenazi Jews. *Fam Cancer*. 2004;3(3-4):241-8.

- Lambowitz AM, Zimmerly S. 2011. Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb Perspect Biol* 3.

- Lander ES, Linton LM, Birren B, Nusbaum C, Zody, MC, Baldwin J, Devon K, Dewar K, Doyle M, Fitz-Hugh W et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.

- Lee E, Iskow R, Yang L, Gokcumen O, Haseley P et al. 2012. Landscape of somatic retrotransposition in human cancers. *Science*. 337(6097):967-71.

- Lehrman MA, Goldstein JL, Russell DW, Brown MS. 1987. Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. *Cell*. 48(5):827-35.

- Lerat E, Capy P. 1999. Retrotransposons and retroviruses: analysis of the envelope gene. *Mol Biol Evol.* 16(9):1198-207.
- Lerman MI, Thayer RE, Singer MF. Kpn I family of long interspersed repeated DNA sequences in primates: polymorphism of family members and evidence for transcription. *Proc Natl Acad Sci U S A.* 80(13):3966-70.
- Léveillé F, Blom E, Medhurst AL, Bier P, Laghmani el H, Johnson M et al. The Fanconi anemia gene product FANCF is a flexible adaptor protein. *J Biol Chem.* 279(38):39421-30.
- Li J, Holzschu DL, Sugiyama T. 2013. PCNA is efficiently loaded on the DNA recombination intermediate to modulate polymerase δ , η , and ζ activities. *Proc Natl Acad Sci USA.* 110(19):7672-7.
- Li X, Zhang J, Jia R, Cheng V, Xu X, et al. 2013. The MOV10 helicase inhibits LINE-1 mobility. *J. Biol. Chem.* 288:21148–60.
- Li X, Stith CM, Burgers PM, Heyer WD. 2009. PCNA is required for initiation of recombination-associated DNA synthesis by DNA polymerase delta. *Mol Cell.* 36(4):704-13.
- Loreto E, Carareto C, Capy P. 2008. Revisiting horizontal transfer of transposable elements in *Drosophila*. *Heredity.* 100:545–554.
- Luan DD, Korman MH, Jakubczak JL, Eickbush TH. 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* 72:595–605.
- Machida YJ, Machida Y, Chen Y, Gurtan AM, Kupfer GM, D'Andrea AD, Dutta A. 2006. UBE2T is the E2 in

the Fanconi anemia pathway and undergoes negative autoregulation. *Mol Cell*. 23(4):589-96.

- MacMillan ML, Wagner JE. 2010. Haematopoietic cell transplantation for Fanconi anaemia - when and how? *Br J Haematol*. 149:14-21.

- Magdalena N, Pilonetto DV, Bitencourt MA, et al. 2005. Frequency of Fanconi anemia in Brazil and efficacy of screening for the FANCA 3788-3790del mutation. *Braz J Med Biol Res*. 38(5):669-673.

- Malik HS, Eickbush TH. 2000. NeSL-1, an ancient lineage of site-specific non-LTR retrotransposons from *Caenorhabditis elegans*. *Genetics*. 154(1):193-203.

- Malik HS, Eickbush TH. 2001. Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. *Genome Res*. 11(7):1187-97.

- Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature*. 424(6944):99-103.

- Martín F, Marañón C, Olivares M, Alonso C, López MC. 1995. Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from *Trypanosoma cruzi*: homology of the first ORF with the ape family of DNA repair enzymes. *J Mol Biol*. 247(1):49-59.

- Martín F, Olivares M, López MC, Alonso C. 1996. Do non-long terminal repeat retrotransposons have nuclease activity? *Trends Biochem Sci*. 21(8):283-5.

- Martin SL. 1991. Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. *Mol. Cell. Biol*. 11:4804-7.

- Martin SL, Bushman FD. 2001. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol. Cell. Biol.* 21:467-75.
- Martin SL, Branciforte D, Keller D, Bain DL. 2003. Trimeric structure for an essential protein in L1 retrotransposition. *Proc Natl Acad Sci U S A.* 2003 Nov 25;100(24):13815-20.
- Martin SL. 2006. The ORF1 protein encoded by LINE-1: structure and function during L1 retrotransposition. *J Biomed Biotechnol.* 2006(1):45621.
- Mathews MB, Bernstein RM, Franza BR Jr, Garrels JI. 1984. Identity of the proliferating cell nuclear antigen and cyclin. *Nature.* 309(5966):374-6.
- Mathias SL, Scott AF, Kazazian HH Jr, Boeke JD, Gabriel A. 1991. Reverse transcriptase encoded by a human transposable element. *Science.* 254(5039):1808-10.
- McDonald JF. 1993. Evolution and consequences of transposable elements. *Curr Opin Genet Dev.* 3(6):855-64. Review.
- McClintock B. 1950. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA* 36:344-55.
- McMillan JP, Singer MF. 1993. Translation of the human LINE-1 element, L1Hs. *Proc Natl Acad Sci U S A.* 1993 Dec 15;90(24):11533-7.
- Meetei AR, Levitus M, Xue Y, Medhurst AL, Zwaan M, Ling C et al. 2004. X-linked inheritance of Fanconi anemia complementation group B. *Nat Genet.* 36(11):1219-24.
- Meetei AR, Medhurst AL, Ling C, Xue Y, Singh TR, Bier P et al. 2005. A human ortholog of archaeal DNA

repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet.* 37(9):958-63.

- Meischl C, Boer M, Ahlin A, Roos D. 2000. A new exon created by intronic insertion of a rearranged LINE-1 element as the cause of chronic granulomatous disease. *Eur. J. Haematol.* 60, 349-350.

- Mi J, Qiao F, Wilson JB, High AA, Schroeder MJ, Stukenberg PT, Moss A, Shabanowitz J, Hunt DF, Jones NJ, Kupfer GM. 2004. FANCG is phosphorylated at serines 383 and 387 during mitosis. *Mol Cell Biol.* 24(19):8576-85.

- Miciak A, Keen A, Jadayel D, Bunday S. 1992. Multiple mutation in an extended Duchenne muscular dystrophy family. *J Med Genet.* 29(2):123-6.

- Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, et al. 1992. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* 52:643-45.

- Miller WJ, McDonald JF, Pinsky W. 1997. Molecular domestication of mobile elements. *Genetica.* 1997;100(1-3):261-70.

- Mills RE, Bennett EA, Iskow RC, Devine SE. 2007. Which transposable elements are active in the human genome? *Trends Genet.* 23(4):183-91.

- Minakami R, Kurose K, Etoh K, Furuhashi Y, Hattori M, Sakaki Y. 1992. Identification of an internal cis-element essential for the human L1 transcription and a nuclear factor(s) binding to the element. *Nucleic Acids Res.* 20:3139-45.

- Mizuuchi K. 1992. Transpositional recombination: mechanistic insights from studies of mu and other elements. *Annu Rev Biochem.* 61:1011-51.

- Moldovan GL, Pfander B, Jentsch S. 2007. PCNA, the maestro of the replication fork. *Cell*. 129(4):665-79
- Moldovan GL, D'Andrea AD. 2009. How the fanconi anemia pathway guards the genome. *Annu Rev Genet*. 43:223-49.
- Molina-Estevez FJ, Lozano ML, Navarro S, Torres Y, Grabundzija I, Ivics Z, Samper E, Bueren JA, Guenechea G. 2013. Brief report: impaired cell reprogramming in nonhomologous end joining deficient cells. *Stem Cells*. 31(8):1726-30.
- Montes de Oca R, Andreassen PR, Margossian SP, Gregory RC et al. 2005. Regulated interaction of the Fanconi anemia protein, FANCD2, with chromatin. *Blood*. 105(3):1003-9.
- Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr. 1996. High frequency retrotransposition in cultured mammalian cells. *Cell* 87:917-27.
- Moran JV, DeBerardinis RJ, Kazazian HH Jr. 1999. Exon shuffling by L1 retrotransposition. *Science* 283:1530-34.
- Morgan LV. 1926. Correlation between Shape and Behavior of a Chromosome. *Proc Natl Acad Sci U S A*. 12(3):180-1.
- Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, et al. 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat. Genet*. 31:159-65.
- Morrish TA, Garcia-Perez JL, Stamato TD, Taccioli GE, Sekiguchi J, Moran JV. 2007. Endonucleaseindependent LINE-1 retrotransposition at mammalian telomeres. *Nature* 446:208-12.
- Müller LU, Milsom MD, Harris CE, Vyas R, Brumme KM, Parmar K, Moreau LA, Schambach A, Park IH,

- London WB, Strait K, Schlaeger T, Devine AL, Grassman E, D'Andrea A, Daley GQ, Williams DA. 2012. Overcoming reprogramming resistance of Fanconi anemia cells. *Blood*. 119(23):5449-57.
- Muñoz-Lopez M, Medina PP, Garcia-Perez JL. 2013. Wiping DNA methylation: wip1 regulates genomic fluidity on cancer. *Cancer Cell*. 24(4):405-7.
 - Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. 2005. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 435:903–10.
 - Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, et al. 2010. L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 468:443–46.
 - Muranyi W, Darai G, Flügel RM. 1988. Nucleotide sequence of a human retrotransposon region encoding a reverse transcriptase-like protein. *Nucleic Acids Res*. 16(15):7724.
 - Nagaraju G, Scully R. 2007. Minding the gap: the underground functions of BRCA1 and BRCA2 at stalled replication forks. *DNA Repair (Amst)*. 6(7):1018-31.
 - Narita N, Nishio H, Kitoh Y, Ishikawa Y, Ishikawa Y, Minami R et al. 1993. Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. *J Clin Invest*. 91(5):1862-7.
 - Newman TL, Tuzun E, Morrison VA, Hayden KE, Ventura M et al. A genome-wide survey of structural variation between human and chimpanzee. *Genome Res*. 15(10):1344-56.
 - Niedernhofer LJ, Odijk H, Budzowska M, van Drunen E et al. 2004. The structure-specific endonuclease

Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. *Mol Cell Biol.* 24(13):5776-87.

- Niedernhofer LJ, Lalai AS, Hoeijmakers JH. 2005. Fanconi anemia (cross)linked to DNA repair. *Cell.*123(7):1191-8.

- Nookala RK, Hussain S, Pellegrini L. 2007. Insights into Fanconi Anaemia from the structure of human FANCE. *Nucleic Acids Res.* 35(5):1638-48.

- Nouaud D, Anxolabéhère D. 1997. P element domestication: a stationary truncated P element may encode a 66-kDa repressor-like protein in the *Drosophila montium* species subgroup. *Mol Biol Evol.* 14(11):1132-44.

- Okada N. 1991. SINEs. *Curr Opin Genet Dev.* 1(4):498-504

- Oku T, Ikeda S, Sasaki H, Fukuda K, Morioka H et al. 1998. Functional sites of human PCNA which interact with p21 (Cip1/Waf1), DNA polymerase delta and replication factor C. *Genes Cells.* 3(6):357-69.

- Onozawa M, Zhang Z, Kim YJ, Goldberg L, Varga T, Bergsagel PL. 2014. Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. *Proc Natl Acad Sci U S A.* 111(21):7729-34.

- Orgel LE, Crick FH. 1980. Selfish DNA: the ultimate parasite. *Nature.* 284(5757):604-7.

- Ostertag EM, Kazazian HH Jr. 2001. Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. *Genome Res.* 11(12):2059-65.

- Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, et al. 2002. A mouse model of human L1 retrotransposition. *Nat. Genet.* 32:655-60.

- Pace P, Mosedale G, Hodskinson MR, Rosado IV, Sivasubramaniam M, Patel KJ. 2010. Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. *Science*. 329(5988):219-23.
- Pichierra P, Rosselli F. 2004. Fanconi anemia proteins and the S phase checkpoint. *Cell Cycle*. 3(6):698-700.
- Pickeral OK, Makałowski W, Boguski MS, Boeke JD. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res*. 10(4):411-5.
- Piguet V, Wan L, Borel C, Mangasarian A, Demaurex N, Thomas G, Trono D. 2000. HIV-1 Nef protein binds to the cellular protein PACS-1 to downregulate class I major histocompatibility complexes. *Nat Cell Biol*. 2(3):163-7.
- Piskareva O, Schmatchenko V. 2006. DNA polymerization by the reverse transcriptase of the human L1 retrotransposon on its own template in vitro. *FEBS Lett*. 580(2):661-8.
- Pochart P, Agoutin B, Rousset S, Chanet R, Doroszkiewicz V, Heyman T. 1993. Biochemical and electron microscope analyses of the DNA reverse transcripts present in the virus-like particles of the yeast transposon Ty1. Identification of a second origin of Ty1DNA plus strand synthesis. *Nucleic Acids Res*. 21(15):3513-20.
- Prelich G, Stillman B. 1988. Coordinated leading and lagging strand synthesis during SV40 DNA replication in vitro requires PCNA. *Cell*. 53(1):117-26.
- Pritham EJ, Putliwala T, Feschotte C. 2007. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene*. 390(1-2):3-17.

- Puck TT, Cieciura SJ, Robinson A. 1958. Genetics of somatic mammalian cells. III. Long-term cultivation of euploid cells from human and animal subjects. *J Exp Med.* 108(6):945-56.
- Purcell DF, Broscius CM, Vanin EF, Buckler CE, Nienhuis AW, Martin MA. 1996. An array of murine leukemia virus-related elements is transmitted and expressed in a primate recipient of retroviral gene transfer. *J Virol.* 70(2):887-97.
- Qiao F, Mi J, Wilson JB, Zhi G, Bucheimer NR, Jones NJ, Kupfer GM. 2004. Phosphorylation of fanconi anemia (FA) complementation group G protein, FANCG, at serine 7 is important for function of the FA pathway. *J Biol Chem.* 279(44):46035-45.
- Rachidi M, Lopes C, Benichou JC, Hellio R, Maisonhaute C. 2005. Virus-like particle formation in *Drosophila melanogaster* germ cells suggests a complex translational regulation of the retrotransposon cycle and new mechanisms inhibiting transposition. *Cytogenet Genome Res.* 111(1):88-95.
- Räschle M, Knipscheer P, Enoiu M, Angelov T, Sun J et al. 2008. Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell.* 134(6):969-80.
- Reichmann J, Reddington JP, Best D, Read D, Ollinger R, Meehan RR, Adams IR. 2013. The genome-defence gene *Tex19.1* suppresses LINE-1 retrotransposons in the placenta and prevents intra-uterine growth retardation in mice. *Hum Mol Genet.* 22(9):1791-806.
- Rosenberg PS, Socié G, Alter BP, Gluckman E. 2005. Risk of head and neck squamous cell cancer and death in patients with Fanconi anemia who did and did not receive transplants. *Blood.* 105(1):67-73.

- Rosenberg PS, Alter BP, Ebell W. 2008. Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry. *Haematologica*. 93(4):511-7.
- Rosenberg PS, Tamary H, Alter BP. 2011. How high are carrier frequencies of rare recessive syndromes? Contemporary estimates for Fanconi Anemia in the United States and Israel. *Am J Med Genet A*. 155A(8):1877-83.
- Raya A, Rodríguez-Pizà I, Guenechea G, Vassena R, Navarro S, Barrero MJ, Consiglio A, Castellà M, Río P, Sleep E, González F, Tiscornia G, Garreta E, Aasen T, Veiga A, Verma IM, Surrallés J, Bueren J, Izpisúa Belmonte JC. 2009. Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature*. 460(7251):53–59.
- Reichmann J, Crichton JH, Madej MJ, Taggart M, Gautier P, et al. 2012. Microarray analysis of LTR retrotransposon silencing identifies Hdac1 as a regulator of retrotransposon expression in mouse embryonic stem cells. *PLoS Comput. Biol.* 8:e1002486.
- Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV. 2014. APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. *eLife* 3:e02008.
- Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, et al. 2010. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* 463:237–40.
- Sakaki Y. 1987. RNA-mediated (?) dispersion of the L1 family of long interspersed repetitive DNA in mammalian genomes. *Mol Biol Med*. 4(4):193-7.

- Sato K, Toda K, Ishiai M, Takata M, Kurumizaka H. 2012. DNA robustly stimulates FANCD2 monoubiquitylation in the complex with FANCI. *Nucleic Acids Res.* 40(10):4553-61.
- Sayah DM, Sokolskaja E, Berthouex L, Luban J. 2004. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature.* 430(6999):569-73.
- Schwahn U, Lenzner S, Dong J, Feil S, Hinzmann B, van Duijnhoven G et al. 1998. Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nat Genet.* 19(4):327-32.
- Seleme MC, Vetter MR, Cordaux R, Bastone L, Batzer MA, Kazazian HH Jr. 2006. Extensive individual variation in L1 retrotransposition capability contributes to human genetic diversity. *Proc Natl Acad Sci U S A.* 103(17):6611-6.
- Sen SK, Huang CT, Han K, Batzer MA. 2007. Endonuclease-independent insertion provides an alternative pathway for L1 retrotransposition in the human genome. *Nucleic Acids Res.* 35:3741-51.
- Shafit-Zagardo B, Maio JJ, Brown FL. 1982. KpnI families of long, interspersed repetitive DNAs in human and other primate genomes. *Nucleic Acids Res.* 10(10):3175-93.
- Sharma NM, Kochenova OV, Shcherbakova PV. 2011. The non-canonical protein binding site at the monomer-monomer interface of yeast proliferating cell nuclear antigen (PCNA) regulates the Rev1-PCNA interaction and Pol ζ /Rev1-dependent translesion DNA synthesis. *J Biol Chem.* 286(38):33557-66.
- She X, Cheng Z, Zöllner S, Church DM, Eichler EE. 2008. Mouse segmental duplication and copy number variation. *Nat Genet.* 40(7):909-14.

- Shimamura A, Alter BP . 2010. Pathophysiology and management of inherited bone marrow failure syndromes. *Blood Rev.* 2010 May;24(3):101-22.
- Shukla R, Upton KR, Munoz-Lopez M, Gerhardt DJ, Fisher ME, et al. 2013. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell* 153:101–11.
- Skowronski J, Singer MF. 1985. Expression of a cytoplasmic LINE-1 transcript is regulated in a human teratocarcinoma cell line. *Proc Natl Acad Sci U S A.* 82(18):6050-4.
- Singer MF. 1982. SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell.* 28(3):433-4.
- Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH. 2010. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends Neurosci.* 33:345–54.
- Smogorzewska A, Desetty R, Saito TT, Schlabach M, Lach FP et al. 2010. A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Mol Cell.* 39(1):36-47.
- Smogorzewska A, Matsuoka S, Vinciguerra P, McDonald ER 3rd, Hurov KE et al. 2007. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell.* 129(2):289-301.
- Solyom S, Ewing AD, Rahrman EP, Doucet T, Nelson HH, et al. 2012. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res.* 22:2328–38.
- Somyajit K, Subramanya S, Nagaraju G. 2010. RAD51C: a novel cancer susceptibility gene is linked

to Fanconi anemia and breast cancer. *Carcinogenesis*.31(12):2031-8.

- Somyajit K, Basavaraju S, Scully R, Nagaraju G. 2013. ATM- and ATR-mediated phosphorylation of XRCC3 regulates DNA double-strand break-induced checkpoint activation and repair. *Mol Cell Biol*. 33(9):1830-44.

- Song SU, Gerasimova T, Kurkulos M, Boeke JD, Corces VG. 1994. An env-like protein encoded by a *Drosophila* retroelement: evidence that gypsy is an infectious retrovirus. *Genes Dev*. 8:2046-2057.

- Speek M. 2001. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol. Cell. Biol*. 21:1973–85.

- Spradling AC, Bellen HJ, Hoskins RA. 2011. *Drosophila* P elements preferentially transpose to replication origins. *Proc Natl Acad Sci U S A*. 108(38):15948-53.

- Stoepker C, Hain K, Schuster B, Hilhorst-Hofstee Y, Roimans MA et al. 2011. SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nat Genet*. 43(2):138-41.

- Suhasini AN, Rawtani NA, Wu Y, Sommers JA, Sharma S et al. 2011. Interaction between the helicases genetically linked to Fanconi anemia group J and Bloom's syndrome. *EMBO J*. 30(4):692-705.

- Sung P, Robberson DL. 1995. DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell*. 82(3):453-61.

- Sverdlov ED. 1998. Perpetually mobile footprints of ancient infections in human genome. *FEBS Lett*. 428(1-2):1-6.

- Swergold GD. 1990. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* 10:6718–29.
- Swift M, Zimmerman D, McDonough ER. 1971. Squamous cell carcinomas in Fanconi's anemia. *JAMA.* 216(2):325-6.
- Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ et al. 2002. Human L1 retrotransposition is associated with genetic instability in vivo. *Cell* 110:327–38.
- Takai D, Yagi Y, Habib N, Sugimura T, Ushijima T. 2000. Hypomethylation of LINE1 retrotransposon in human hepatocellular carcinomas, but not in surrounding liver cirrhosis. *Jpn J Clin Oncol.* 30(7):306-9.
- Takasu M, Hayashi R, Maruya E, Ota M, Imura K, et al. 2007. Deletion of entire HLA-A gene accompanied by an insertion of a retrotransposon. *Tissue Antigens* 70:144–50.
- Tamary H, Fromovich-Amit Y, Shalmon L, Zaizov R, Yaniv I et al. 2000. Molecular characterization of four novel mutations causing factor VII deficiency. *Hematol J.* 1(6):382-9.
- Taniguchi T, Garcia-Higuera I, Xu B, Andreassen PR, Gregory RC et al. 2002. Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. *Cell.* 109(4):459-72.
- Taylor MS, Lacava J, Mita P, Molloy KR, Huang CR, et al. 2013. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell* 155:1034–48.
- Tebbs RS, Hinz JM, Yamada NA, Wilson JB, Salazar EP, Thomas CB, Jones IM, Jones NJ, Thompson LH. 2005. New insights into the Fanconi anemia pathway

from an isogenic FancG hamster CHO mutant. DNA Repair (Amst). 4(1):11-22.

- Thayer RE, Singer MF. 1983. Interruption of an alpha-satellite array by a short member of the KpnI family of interspersed, highly repeated monkey DNA sequences. Mol Cell Biol. 3(6):967-73.

- Thompson LH. 2002. Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: the molecular choreography. Mutat Res. 751(2):158-246.

- Thompson LH, Hinz JM. 2009. Cellular and molecular consequences of defective Fanconi anemia proteins in replication-coupled DNA repair: mechanistic insights. Mutat Res. 668(1-2):54-72.

- Thompson LH. 2012. Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: the molecular choreography. Mutat Res. 751(2):158-246.

- Ton-Hoang B, Guynet C, Ronning DR, Cointin-Marty B, Dyda F, Chandler M. 2005. Transposition of ISHp608, member of an unusual family of bacterial insertion sequences. EMBO J. 24(18):3325-38.

- van den Hurk JA, Meij IC, Seleme MC, Kano H, Nikopoulos K, et al. 2007. L1 retrotransposition can occur early in human embryonic development. Hum. Mol. Genet. 16:1587-92.

- van der Lelij P, Oostra AB, Rooimans MA, Joenje H, de Winter JP. 2010. Diagnostic overlap between Fanconi anemia and the Cohesinopathies: Roberts Syndrome and Warsaw Breakage Syndrome. Anemia. 2010:565268.

- Smith CH. 2014. Wallace, Darwin and Ternate 1858. Notes Rec R Soc Lond. 2014 Jun 20. 68(2):165-70.

- Wahba AJ, Gardner RS, Basilio C, Miller RS, Speyer JF, Lengyel P. Synthetic polynucleotides and the amino acid code. VIII. Proc Natl Acad Sci U S A. 49:116-22.
- Waisfisz Q, Morgan NV, Savino M, de Winter JP, van Berkel CG, Hoatlin ME, Ianzano L, Gibson RA, Arwert F, Savoia A, Mathew CG, Pronk JC, Joenje H. 1999. Spontaneous functional correction of homozygous fanconi anaemia alleles reveals novel mechanistic basis for reverse mosaicism. Nat Genet. 22(4):379-83.
- Wang W. 2007. Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. Nat Rev Genet. 8(10):735-48.
- Wang X, Kennedy RD, Ray K, Stuckert P, Ellenberger T, D'Andrea AD. 2007. Chk1-mediated phosphorylation of FANCD1 is required for the Fanconi anemia/BRCA pathway. Mol Cell Biol. 27(8):3098-108.
- Wang C, Lambert MW. 2010. The Fanconi anemia protein, FANCD1, binds to the ERCC1-XPF endonuclease via its tetratricopeptide repeats and the central domain of ERCC1. Biochemistry. 49(26):5560-9.
- Warbrick, E. 1998. PCNA binding through a conserved motif. Bioessays 20, 195-199.
- Watson JD, Crick FH. 1953. The structure of DNA. Cold Spring Harb Symp Quant Biol. 18:123-31.
- Wei W, Morrish TA, Alisch RS, Moran JV. 2000. A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. Anal Biochem. 284(2):435-8.
- Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, et al. 2001. Human L1 retrotransposition: cis preference versus trans complementation. Mol. Cell. Biol. 21:1429-39.

- Wessler SR, Bureau TE, White SE. 1995. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Curr Opin Genet Dev.* 5(6):814-21.
- Wilhelm M, Wilhelm FX. 2001. Reverse transcription of retroviruses and LTR retrotransposons. *Cell Mol Life Sci.* 58(9):1246-62.
- Wilson JB, Yamamoto K, Marriott AS, Hussain S, Sung P et al. 2008. FANCG promotes formation of a newly identified protein complex containing BRCA2, FANCD2 and XRCC3. *Oncogene.* 27(26):3641-52.
- Wilson JB, Blom E, Cunningham R, Xiao Y, Kupfer GM, Jones NJ. 2010. Several tetratricopeptide repeat (TPR) motifs of FANCG are required for assembly of the BRCA2/D1-D2-G-X3 complex, FANCD2 monoubiquitylation and phleomycin resistance. *Mutat Res.* 689(1-2):12-20.
- Wissing S, Muñoz-Lopez M, Macia A, Yang Z, Montano M et al. 2012. Reprogramming somatic cells into iPS cells activates LINE-1 retroelement mobility. *Hum Mol Genet.* 21(1):208-18.
- Xia B, Dorsman JC, Ameziane N, de Vries Y, Rooimans MA et al. 2007. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat Genet.* 39(2):159-61
- Xing J , Wang H, Belancio VP, Cordaux R, Deininger PL, Batzer MA. 2006. Emergence of primate genes by retrotransposon-mediated sequence transduction. *Proc Natl Acad Sci U S A.* 103(47):17608-13.
- Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W et al. 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat Biotechnol.* 29(8):735-41.

- Yamamoto K, Ishiai M, Matsushita N, Arakawa H, Lamerdin JE et al. 2003. Fanconi anemia FANCG protein in mitigating radiation- and enzyme-induced DNA double-strand breaks by homologous recombination in vertebrate cells. *Mol Cell Biol.* 23(15):5421-30.
- Yamashita T, Barber DL, Zhu Y, Wu N, D'Andrea AD. 1994. The Fanconi anemia polypeptide FACC is localized to the cytoplasm. *Proc Natl Acad Sci U S A.* 91(14):6712-6.
- Yang N, Zhang L, Zhang Y, Kazazian HH Jr. 2003. An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Res.* 31:4929-40.
- Yang N, Kazazian HH Jr. 2006. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat. Struct. Mol. Biol.* 13:763-71.
- Yao J, Truong DM, Lambowitz AM. 2013. Genetic and biochemical assays reveal a key role for replication restart proteins in group II intron retrohoming. *PLoS Genet.* 9(4):e1003469.
- Yi JM, Kim TH, Huh JW, Park KS, Jang SB, Kim HM, Kim HS. 2004. Human endogenous retroviral elements belonging to the HERV-S family from human tissues, cancer cells, and primates: expression, structure, phylogeny and evolution. *Gene.* 342(2):283-92.
- Yoshida K, Nakamura A, Yazaki M, Ikeda S, Takeda S. 1998. Insertional mutation by transposable element, L1, in the DMD gene results in X-linked dilated cardiomyopathy. *Hum Mol Genet.* 7(7):1129-32.
- Yu F, Zingler N, Schumann G, Stratling WH. 2001. Methyl-CpG-binding protein 2 represses LINE-1

expression and retrotransposition but not Alu transcription. *Nucleic Acids Res.* 29:4493–501.

- Zhang P, Mo JY, Perez A, Leon A, Liu L, Mazloum N, Xu H, Lee MY. 1999. Direct interaction of proliferating cell nuclear antigen with the p125 catalytic subunit of mammalian DNA polymerase delta. *J Biol Chem.* 274(38):26647-53.

- Zhao K, Du J, Han X, Goodier JL, Li P, Zhou X, Wei W, Evans SL, Li L, Zhang W, Cheung LE, Wang G, Kazazian HH Jr, Yu XF. 2013. Modulation of LINE-1 and Alu/SVA Retrotransposition by Aicardi-Goutières Syndrome-Related SAMHD1. *Cell Rep.* 4(6):1108-15.

- Zhi G, Wilson JB, Chen X, Krause DS, Xiao Y, Jones NJ, Kupfer GM. 2009. Fanconi anemia complementation group FANCD2 protein serine 331 phosphorylation is important for fanconi anemia pathway function and BRCA2 interaction. *Cancer Res.* 69(22):8775-83.