



Universitat Autònoma de Barcelona

Facultat de Biociències

Departament de Genètica i Microbiologia

Unitat de Genètica

Secuenciación del exoma en anemia de Fanconi: del diagnóstico al descubrimiento de un nuevo gen

Memoria presentada por **Massimo Bogliolo** para optar al grado de Doctor en Genética por la
Universitat Autònoma de Barcelona

Director de tesis

Autor

Dr. Jordi Surrallés Calonge

Massimo Bogliolo

Catedrático de Genética, UAB

**Esta tesis es dedicada a todos los pacientes de
Anemia de Fanconi y a sus familias**

Contenido

ABREVIATURAS.....	9
I. PROLOGO	13
II. INTRODUCCIÓN.....	17
II.1. La Anemia de Fanconi.....	17
II.1.1. Anomalías del desarrollo.....	17
II.1.2. Defectos de las células madre	18
II.1.3. Neoplasias	19
II.1.4. Mosaicismo somático.....	20
II.1.5. Fenotipo de las células FA	21
II.1.6. FA y aldehídos endógenos	22
II.2. Clonación de los genes FA.....	24
II.3. Genes FA, FA-like y asociados a FA	32
II.4. Biología molecular de la FA.....	33
II.4.1. FAN1	38
II.5. Reparación por Recombinación Homóloga y FA	39
II.6. FA y la susceptibilidad al cáncer familiar de mama y ovario.....	43
II.6.1. Mutaciones bialélicas en genes FBOC: fenotipos celulares y clínicos ..	45
II.7. La secuenciación del DNA de nueva generación: secuenciación del exoma	46
II.7.1. WES: Captura y amplificación de las regiones codificantes.....	47
II.7.2. Secuenciación de las librerías	50
II.7.3. Otras tecnologías para secuenciación de librerías.....	54
II.7.4. NGS y diagnóstico de la FA.....	56
III. OBJETIVOS.....	63

IV.	MATERIALES Y MÉTODOS	67
	IV.1. Pacientes y controles.....	67
	IV.2. Ensayo de fragilidad cromosómica	68
	IV.3. WES y análisis de las mutaciones en pacientes FA.	69
	IV.4. MLPA	70
	IV.5. Líneas celulares.....	71
	IV.6. Análisis del bloqueo del ciclo celular en G2 y citometría de flujo.....	72
	IV.7. Clonaje específico de alelos.....	73
	IV.8. Ensayo de sensibilidad a los ICL y UVC en células adherentes.....	73
	IV.9. Ensayo de la sensibilidad a los ICL y UVC en LCLs	74
	IV.10. Producción de las líneas revertidas FA104R and 1333R.	75
	IV.11. Mutagénesis sitio dirigida	76
	IV.12. Vectores y producción de partículas virales	77
	IV.13. Selección de células por fluorescencia	78
	IV.14. Ensayo de síntesis de DNA no programada	79
	IV.15. Irradiación local con UVC y reparación del daño al DNA.....	80
	IV.16. Western Blot, co-IP y fraccionamiento celular	81
	IV.17. Evaluación del desenganche (<i>unhooking</i>) por ensayo cometa	82
	IV.18. Purificación de XPF-ERCC1 y ensayos de reparación <i>in vitro</i>	83
V.	RESULTADOS	89
	V.1. Subtipaje y caracterización molecular de los pacientes FA mediante método WES + MLPA.....	89
	V.1.1. Análisis de la cobertura de los genes FA en el análisis por WES.....	89
	V.1.2. Análisis por WES y MLPA de 49 pacientes FA.....	91
	V.1.3. Análisis por cobertura de las grandes deleciones en FANCA	93

V.2. Mutaciones bialélicas en <i>ERCC4</i> causan FA	96
V.3. Mutaciones en <i>ERCC4</i> que causan FA afectan marginalmente la NER....	104
V.4. Mutaciones en <i>ERCC4</i> que causan FA inactivan la reparación de los ICL107	
V.5. Las proteínas XPF mutadas en FA relocalizan de forma normal al núcleo y al DNA dañado.	109
V.6. Escisión, actividad nucleasa y “unhooking” de los mutantes XPF	111
V.7. <i>ERCC4/FANCO</i> como posible gen FBOC	112
VI. DISCUSIÓN.....	119
VI.1. Caracterización genética mediante WES + MLPA de 49 pacientes FA... 119	
VI.1.1. Espectro mutacional de los pacientes FA-A analizados	124
VI.1.2. Espectro mutacional de los pacientes FA-A españoles	126
VI.1.3. Espectro mutacional de los pacientes no FA-A analizados	129
VI.2. Identificación de <i>ERCC4</i> como gen FA.....	133
VI.3. <i>ERCC4</i> y FBOC	140
VII. COCLUSIONES	145
VIII. BIBLIOGRAFÍA:.....	149
IX. ANEXOS.....	175

ABREVIATURAS

6-4PP: fotoproductos pirimidina (6-4) pirimidona

8-MOP: 8-Methoxypsoralen

ALL: Leucemia linfocítica aguda

AML: Leucemia Mieloide Aguda

BMF: Fallo de la médula ósea; del inglés, Bone Marrow failure.

bp: pares de bases; del Inglés, base pairs

CGH: Hibridación Genómica Comparada

CisPt: Cisplatino

CNV: Variación del número de copias

CPD: Dímeros de ciclobutano de pirimidinas

CRT: Terminación cíclica reversible

CS: Síndrome de Cockayne

DEB: Diepoxi-butano

DHPLC: Cromatografía líquida desnaturante de alta rendimiento; del inglés, Denaturing High-Performance Liquid Chromatography

DMEM: Medio de cultivo celular Dulbecco's Modified Eagle's Medium

dNTP: desoxirribonucleótido trifosfato

DSB: Roturas de doble cadena del DNA; del inglés, Double Strand Breaks

dsDNA: DNA de cadena doble; del inglés double strand DNA

emPCR: PCR en emulsión.

FA: Anemia de Fanconi

FAAP: Proteínas Asociadas a la Anemia de Fanconi

FACS: Clasificación de células activadas por fluorescencia; del inglés, Fluorescence-Activated Cell Sorting

FBOC: Cáncer de mama y Ovario familiar; del inglés, familial breast and ovary cancer

FCS: Suero Fetal Bovino; del inglés, Fetal Calf Serum

Full-FA: Pacientes con Anemia de Fanconi que no presentan mosaicismo somático

HNSCC: Carcinoma de células escamosas de cabeza y cuello; del inglés, Head and Neck Squamous Cells Carcinoma.

HRR: Reparación por Recombinación Homóloga.

HSC: células madres hematopoyéticas; del inglés, Hematopoyetic Stem Cells

ICL: enlaces entrecruzado entre las dos cadenas del DNA; del inglés, Interstrand crosslinks

IF: inmunofluorescencia

IGV: software de análisis: Integrative Genomics Viewer

InDel: inserciones/deleciones

IP: inmunoprecipitación

IR: radiación ionizante

IRES: Sitio Interno de entrada al Ribosoma; del inglés, Internal ribosome entry site

LCL: Línea celular linfoblastoide inmortalizada

MAF: Prevalencia del Alelo Minoritario, del inglés, Minor Allele Frequency

MEF: Fibroblastos embrionarios de ratón inmortalizados; del inglés, Mouse Embryonic Fibroblasts

MLPA: Amplificación dependiente de ligasa de múltiples sondas; del inglés, Multiplex Ligation-dependent Probe Amplification

MMC: Mitomicina C

MMCT: Transferencia de Cromosomas Mediada por Microcélulas

NER: Reparación por escisión de nucleótidos

NGS: Secuenciación del DNA de nueva Generación; del inglés, Next Generation Sequencing

NHEJ: reparación por unión no homóloga de los extremos; del inglés, Non Homologous End Joining

PB: Sangre periférica; del inglés, Peripheral Blood

RPMI: Medio de Cultivo celular Roswell Park Memorial Institute

SBL: secuenciación por ligación; del inglés, Sequencing by ligation

SCE: Intercambio entre Cromátidas hermanas; del inglés, Sister Chromatid Exchange

sfPCR: PCR en fase solida

SMRT: secuenciación de moléculas únicas de DNA en tiempo real; del inglés, Single Molecule Real-time Sequencing

SNP: Polimorfismo de único nucleótido

SNV: Variantes de único nucleótido; del inglés, Single nucleotide variant

ssDNA: DNA de cadena sencilla; del inglés, Single Strand DNA

TLS: Síntesis de DNA translesión

TM: Momento de la cola; del inglés, Tail Moment

TTD: tricotodistrofia

UBZ: dominios de interacción con la ubiquitina; del inglés, Ubiquitin-binding Zinc finger Domain

UDS: Síntesis de DNA no programada; del inglés, Unscheduled DNA Syntesis

UV: Radiación ultravioleta

UVC: Radiación Ultravioleta de tipo C

WB: Western Blot

WES: Secuenciación del Exoma; del inglés, Whole Exome Sequencing

WGS: Secuenciación del Genoma; del inglés, Whole Genome Sequencing

Wt: tipo natural o silvestre; del inglés, Wild Type

XP: Xeroderma Pigmentosum

"Fatti non foste a viver come bruti ma per seguir virtute e canoscenza"

Dante Alighieri

I. PROLOGO

Debido a su importancia vital, a menudo se asume que el DNA es una molécula fundamentalmente estable. Sin embargo, el DNA es una molécula dinámica que cambia constantemente al reaccionar con agentes endógenos presentes en nuestro cuerpo y con agentes exógenos presentes en nuestro entorno. De hecho, cada una de nuestras células es bombardeada constantemente por agentes que dañan su DNA y amenazan la integridad de su genoma. La formación y la persistencia de las lesiones en el genoma humano pueden conducir a mutagénesis, inestabilidad genética y cáncer (Friedberg et al., 2006a) y, por tanto, las células evolucionaron mecanismos complejos para reparar los daños en el DNA y así mantener la fidelidad del genoma (Wood et al., 2001). La importancia de las vías de reparación del DNA en humanos queda demostrada por la existencia de por lo menos 15 síndromes genéticos asociados con defectos de la reparación del DNA o con las respuestas a los daños en el DNA. En la mayoría de los síndromes hay también una incidencia de cáncer muy elevada y defectos en muchos órganos (Lehmann, 2003). Por ejemplo, la radiación ultravioleta (UV), presente en la luz solar, es un conocido agente físico capaz de dañar el DNA (Lober and Kittler, 1977, Yoon et al., 2000). La Reparación del DNA por Escisión de Nucleótidos (NER) es el principal mecanismo de reparación de estas lesiones antes de que se conviertan en tóxicas y mutagénicas. Cambios en la expresión o en la función de los genes cuyas proteínas realizan la NER, pueden conducir a la persistencia de estas lesiones en el DNA, dejando nuestro genoma vulnerable a la mutagénesis y nosotros al cáncer. La enfermedad genética humana llamada Xeroderma Pigmentosum (XP) es el resultado de mutaciones en algunos de los genes que codifican para las proteínas involucradas en la NER. La XP identificada ya en el siglo diecinueve (Hebra, 1874, Kaposi, 1883), fue el primer desorden que se asoció con un defecto de la reparación del DNA (Cleaver, 1968). Los individuos que sufren de XP son muy sensibles a la luz del sol y tienen una grave

predisposición al cáncer de piel inducido por UV. El sistema NER en estos individuos es defectuoso debido a una mutación en uno de los ocho genes (desde *XPA* hasta *XGP* y *XPV*) involucrados en la NER (Lehmann et al., 2011). El XP es una enfermedad rara autosómica recesiva y las estimaciones realizadas en Europa occidental sugieren aproximadamente 2,3 casos por millón de nacimientos (Kleijer et al., 2008) y en Japón de 1 en 20, 000 (Hirai et al., 2006). El XP por lo tanto es una enfermedad muy poco frecuente y algunos grupos de complementación como por ejemplo el F (XP-F) asoman a poco más de una decena de pacientes en todo el mundo (Hirai et al., 2006). Sin embargo, el estudio de una enfermedad tan rara y la caracterización de los genes mutados en estos pocos pacientes, llevaron al descubrimiento y a la caracterización de una de las rutas más importantes de la reparación del DNA con consecuencias trascendentales también para la población en general. De forma parecida, en los últimos años el estudio de otra enfermedad rara, la Anemia de Fanconi (FA), ha desvelado, aunque todavía no por completo, otro complejo mecanismo de reparación del daño al DNA cuya alteraciones producen inestabilidad genética y predisposición al cáncer tanto en los pacientes como en la población general (Deans and West, 2011). Por lo tanto, los síndromes genéticos humanos asociados con defectos de la reparación del DNA son excelentes modelos para dilucidar los procesos por los cuales los fallos en la reparación del DNA pueden conducir a la inestabilidad genética y a enfermedades humanas (Surralles et al., 2004).

INTRODUCCIÓN

“You just come along with me and have a good time. The Galaxy's a fun place. You'll need to have this fish in your ear.”

Ford Prefect

II. INTRODUCCIÓN

II.1. La Anemia de Fanconi

El médico Guido Fanconi describió por primera vez la FA en 1927 en tres niños que sucumbieron a una anemia asociada con microcefalia, hiperpigmentación de la piel e hipoplasia del testículo. Pronto se dio cuenta de que esta enfermedad afectaba todos los linajes celulares de la sangre y se asociaba con defectos del desarrollo de muchos órganos y a predisposición a desarrollar cáncer (Fanconi, 1927). La FA es una enfermedad clínicamente heterogénea, hereditaria y recesiva que en las personas afectadas generalmente culmina en una insuficiencia de la médula ósea. Desde el descubrimiento del primer gen FA hace 20 años, se han identificado hasta ahora 19 genes asociados con la FA, los cuales incluyen *FANCA*, *FANCB*, *FANCC*, *FANCD1 (BRCA2)*, *FANCD2*, *FANCE*, *FANCF*, *FANCG (XRCC9)*, *FANCI*, *FANCL (BRIP1)*, *FANCL (PHF9)*, *FANCM*, *FANCN (PALB2)*, *FANCO (RAD51C)*, *FANCP (SLX4)*, *FANCO (ERCC4)*, *FANCS (BRCA1)*, *FANCT (UBE2T)* y *FANCR (RAD51)* (Bogliolo et al., 2013, Sawyer et al., 2014, Virts et al., 2015, Rickman et al., 2015, Hira et al., 2015, Wang et al., 2015). La FA afecta aproximadamente a 1 de cada 360.000 nacimientos y los pacientes fallecen de media a los 20 años (Auerbach, 2009). Los afectados por FA exhiben una amplia gama de características clínicas, cada una de penetrancia incompleta, y las principales características clínicas pueden dividirse generalmente en tres categorías: anomalías del desarrollo, defectos de las células madre y neoplasias (Crossan and Patel, 2012).

II.1.1. Anomalías del desarrollo

El fenotipo de malformaciones de los pacientes de FA es muy heterogéneo: los afectados por FA van desde tener un desarrollo prácticamente normal, hasta tener anomalías

de tal gravedad que pueden resultar en aborto espontáneo o letalidad perinatal (Cox et al., 1997, Holden et al., 2006). El 60–70% de los pacientes FA tiene al menos una malformación congénita (Kutler et al., 2003) y las anomalías más observadas en FA son defectos de las extremidades superiores como la hipoplasia o la ausencia completa del radio y las alteraciones de los pulgares. Muy frecuentemente se observan también baja estatura, malformaciones o ausencia de los riñones, defectos del tracto urinario, alteraciones gastrointestinales, defectos cardíacos, anomalías de los ojos, de las orejas con pérdida de la audición y de la piel (hipopigmentación o manchas “café con leche”) (Auerbach, 2009, Joenje and Patel, 2001). Las malformaciones congénitas observadas en FA probablemente representan el resultado final de una apoptosis anormal durante la embriogénesis y es probable que este incremento en la muerte celular sea debido a la incapacidad de las células de reparar los daños en el DNA que se producen normalmente durante el desarrollo (Liu et al., 2003, Castella et al., 2011a).

II.1.2. Defectos de las células madre

En primer lugar, hay el defecto de las células madres hematopoyéticas (*Hematopoietic Stem Cell*; HSC) que resulta en la anemia aplásica, una de las principales causas de morbilidad y mortalidad en los pacientes de FA. Los valores hematológicos suelen ser normales al nacer, pero a lo largo de la vida (y casi siempre en edad pediátrica) los pacientes desarrollan una pancitopenia progresiva (Kutler et al., 2003, Butturini et al., 1994) cuya aparición está asociada a la extensión y la gravedad de las malformaciones del desarrollo (Rosenberg et al., 2004, Tulpule et al., 2010). Estas observaciones sugieren que el feto afecto por FA pueda estar expuesto en el útero a una fuente endógena y estocástica de daños al DNA que producen las anomalías del desarrollo y el agotamiento de las reservas hematopoyéticas embrionarias y que, posteriormente durante el crecimiento, puede promover una progresión más rápida hacia la anemia aplásica dentro de la primera década de vida. Los pacientes de FA también tienen defectos en el acervo de las células madres germinales que acaban diferenciándose en espermatozoides y óvulos en el adulto. Los pacientes varones de FA tienen una fertilidad y un conteo de espermatozoides considerablemente reducido y una mayor prevalencia de espermatozoides con morfología anormal. Los pacientes femeninos

también suelen ser estériles, con un aumento de la prevalencia del fallo ovárico y de la menopausia prematura respecto a la población en general (Auerbach, 2009, Kutler et al., 2003) Contrariamente al fenotipo del fallo medular (*Bone Marrow Failure*; BMF), el fenotipo de esterilidad de los pacientes es también compartido por los modelos murinos de FA. De hecho, todos los modelos de FA de ratón desarrollados hasta hoy muestran el mismo defecto de fertilidad aunque no tengan anemia (Parmar et al., 2009, Crossan et al., 2011, Nabhan et al., 2010). Los ratones mutantes en genes responsables de la FA nacen con muy pocas células germinales en el testículo o en el ovario. Sin embargo los ratones FA pueden procrear, indicando que la ruta FA no es esencial para el desarrollo de las células germinales (Cheng et al., 2000) y que el defecto es debido probablemente a la acumulación de daños en el DNA que induce la apoptosis de las células germinales primordiales.

II.1.3. Neoplasias

Los pacientes FA tienen una gran predisposición a desarrollar cáncer principalmente leucemia mieloide aguda (AML) y tumores sólidos como los carcinomas escamosos de cabeza y cuello (*Head and Neck Squamous Cell Carcinoma*; HNSCC). En total, el 9% de los pacientes de FA llega a desarrollar leucemia, con una media de 14 años de edad al diagnóstico. La prevalencia de AML en la población de FA es en marcado contraste con la de la población no-FA en la cual la mayoría de las leucemias en edad pediátrica son las leucemias linfoides agudas (ALL) (Butturini et al., 1994, Kutler et al., 2003, Alter et al., 2003, Rosenberg et al., 2003, Alter et al., 2005). Las alteraciones genéticas en AML en FA afectan a muchos oncogenes, que han sido implicados previamente en AML no-FA (Quentin et al., 2011). Se puede por lo tanto imaginar que en un embrión FA, a causa de los daños en el DNA que no se pueden reparar correctamente, se producen a la vez el agotamiento de las células madres hematopoyéticas, que produce la anemia aplásica, y también los eventos iniciales en la génesis de un clon leucémico que producirá AML en la primera infancia (Greaves and Wiemels, 2003). Los pacientes diagnosticados en la edad adulta, a menudo no tienen ninguna característica de la enfermedad excepto la inestabilidad genómica y la predisposición a desarrollar tumores sólidos. Esta predisposición es muy significativa, con una incidencia acumulativa de 76% en la quinta década (Rosenberg et al., 2003) y varios

pacientes presentan múltiples tumores primarios independientes. La mayoría de los tumores sólidos son HNSCC, tumores del esófago y de la región ano genital (Alter et al., 2003, van Zeeburg et al., 2008, Rosenberg et al., 2008). Probablemente los pacientes diagnosticados en edad adulta estuvieron expuestos a menos daños al DNA a lo largo del desarrollo embrionario y el fenotipo clínico no se reveló hasta la edad adulta con la formación de tumores sólidos.

II.1.4. Mosaicismo somático

El mosaicismo somático es un fenómeno que afecta entre el 15 y el 25% de los pacientes de FA (Soulier et al., 2005, Lo Ten Foe et al., 1997). Aparece cuando una única HSC revierte espontáneamente la mutación de uno de los dos alelos mutados y, al ser la FA una enfermedad recesiva, revierte también el fenotipo celular convirtiéndose en una célula “normal”. Debido a su ventaja proliferativa respecto a las células que no se han revertido, esta célula se expande de forma clonal y coloniza la médula ósea del paciente (Kwee et al., 1983, Lo Ten Foe et al., 1997). El mosaicismo somático no es un fenómeno exclusivo de la FA, pero siempre se ha observado en tejido hematopoyético donde la alta proliferación celular y la habilidad de las células de colonizar la médula ósea hacen que este fenómeno sea posible (Hirschhorn et al., 1996). El grado de colonización de la médula por parte de las células revertidas puede ser muy variable, desde solo unas pocas células hasta casi la totalidad de la médula. La reversión original de la mutación puede producirse en una célula madre o en una célula más diferenciada de alguno de los linajes hematopoyéticos. Dependiendo del estado de diferenciación de la célula, la reversión involucrará todos los linajes hematopoyéticos o solo parte de ellos. Cuando la reversión se produce a nivel de una HSC los recuentos hematopoyéticos del paciente pueden revertir a niveles normales, en un proceso de “terapia génica natural” (Youssoufian, 1996, Lo Ten Foe et al., 1997). A nivel molecular, los mecanismos que pueden generar la reversión de la mutación son varios: conversión génica, recombinación intragénica o reversión directa de la mutación generando un alelo completamente normal, mientras que la aparición de mutaciones compensatorias en *cis* puede producir una restitución parcial o total de la funcionalidad de la proteína (Gross et al., 2002, Lo Ten Foe et al., 1997).

II.1.5. Fenotipo de las células FA

La sensibilidad a los agentes químicos que producen enlaces entrecruzados (*Inter-strand* crosslinks; ICL) entre las dos cadenas complementarias del DNA es la principal característica de todas las células que derivan de pacientes de FA (Sasaki, 1975). Los ICL son altamente dañinos para la célula porque constituyen un impedimento para el avance de la horquilla de replicación y además, afectando a las dos cadenas del DNA, complican la reparación libre de errores al no ser disponible una cadena intacta de DNA que pueda ser utilizada como molde. Esta hipersensibilidad de las células de FA a agentes como el diepoxibutano (DEB), mitomicina C (MMC), psoralenos fotoactivados (8-MOP) o cisplatino (CisPt) es tan específica que se utiliza para el diagnóstico de la enfermedad. El ensayo tradicional y también el más utilizado para la confirmación del diagnóstico clínico de FA es el test de fragilidad cromosómica. Al tratar con DEB o MMC las células de un paciente de FA, la incapacidad de estas de reparar los ICL hace que en la metafase se puedan observar respecto al control un gran número de roturas de tipo cromatídico que pueden derivar en la formación de estructuras tri- y tetra-radiales como resultado de la mala unión de las extremidades cromosómicas rotas (Figura 1A) (Auerbach, 1988). La alteración del ciclo celular es también una característica de las células FA: como resultado de la acumulación del daño y de la incapacidad de repararlo, las células FA suelen acumularse en la fase G2/M del ciclo celular sobre todo después de tratamientos con agentes que inducen ICL (Dutrillaux et al., 1982, Kubbies et al., 1985) (Figura 1B). Además de la sensibilidad a los ICL, las células FA tienen otras anomalías fenotípicas (Tabla I) que se han estudiado a lo largo de los años en células procedentes de pacientes de FA que pertenecían a diferentes grupos de complementación y, por consecuencia, no queda claro si estas alteraciones son comunes a todas las células FA o solo a un subconjunto. La mayoría de las anomalías descritas son probablemente fenómenos accesorios y no se relacionan directamente con el principal defecto celular de cada grupo de complementación (Grompe and D'Andrea, 2003, D'Andrea and Grompe, 1997). Todas estas características probablemente explican una parte del fenotipo de los pacientes (Tabla I), pero la verdadera comprensión de la enfermedad solo puede provenir de un entendimiento del papel de las proteínas FA a nivel molecular (D'Andrea and Grompe, 1997).

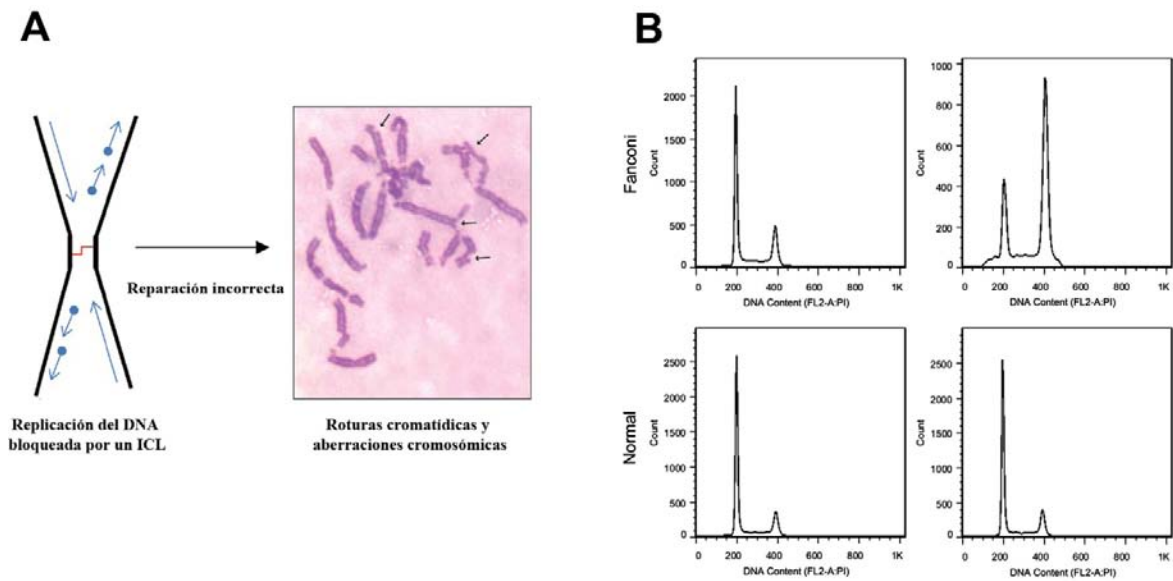


Figura 1: (A) Deficiencias en genes de la FA resultan en un aumento de la prevalencia de las aberraciones cromosómicas (flechas) después de tratamiento con agentes que producen ICL. (B) Análisis del ciclo celular mediante citometría de flujo de un paciente de FA (arriba) y de un individuo normal, sin (NT) y tratados con MMC.

Tabla I: Anomalías celulares en FA:

Fenotipos celulares comunes a todo los grupos de complementación	Referencias
Sensibilidad a los ICLs	(Auerbach and Wolman, 1976)
Roturas cromosómicas espontaneas y inducida por ICLs	(Schroeder, 1964, Auerbach and Wolman, 1978)
Prolongación de la fase G2 del ciclo celular	(Kaiser and S.A., 1982)
Fenotipos celulares estudiados solo en algunos de los grupos de complementación	Referencias
Crecimiento perjudicado por altos niveles de O ₂	(Joenje et al., 1981)
Sobreproducción de radicales del O ₂	(Korkina et al., 1992)
Peores defensas contra el estrés oxidativo	(Gille et al., 1987)
Deficiencias en la superóxido dismutasa	(Joenje et al., 1979)
Sensibilidad a la radiación ionizante	(Bigelow et al., 1979, Taniguchi et al., 2002b)
Sobreproducción de tumor necrosis factor- α (TNF- α)	(Rosselli et al., 1992)
Acumulación de bases oxidadas en el DNA	(Takeuchi and Morimoto, 1993)
Incremento de las mutaciones por deleción	(Papadopoulo et al., 1990)
Incremento de la apoptosis	(Willingale-Theune et al., 1989)
Defectos en los puntos de control del ciclo celular en respuesta a daño al DNA	(Sala-Trepat et al., 2000, Castillo et al., 2011, Taniguchi et al., 2002b, Nakanishi et al., 2002)

II.1.6. FA y aldehídos endógenos

Los productos secundarios del metabolismo embrionario y/o materno siempre se han postulado como las fuentes más probables del daño al DNA que afecta al desarrollo de los embriones FA pero las especies químicas responsables en concreto quedaban desconocidas.

Recientemente los aldehídos han centrado el atención de la investigación de la FA como una de las especies químicas candidatas a ser unas de las causas más importantes de daño endógeno en los embriones FA. El metabolismo normalmente genera aldehídos, conocidos por ser carcinógenos comunes que, sin catabolismo enzimático, pueden acumularse y causar daños al DNA (O'Brien et al., 2005). Una serie de estudios pioneros en ratones han demostrado que los genes Fanconi contrarrestan la genotoxicidad inducida por el acetaldehído y que el daño al DNA generado por los aldehídos podría contribuir de forma muy importante al desarrollo anormal, al fallo hematopoyético y a la predisposición al cáncer de los pacientes FA (Langevin et al., 2011). ALDH2, que media el metabolismo del acetaldehído, es esencial para el desarrollo embrionario de ratones *Fancd2*^{-/-} y, de manera consistente, ALDH5 y la vía FA median letalidad sintética también en las células de pollo DT40 (Langevin et al., 2011, Rosado et al., 2011). Los embriones *Aldh2*^{-/-} *Fancd2*^{-/-} no son viables si la madre carece de ALDH2 funcional. Cuando la madre es *Aldh2*^{+/-}, los ratones *Aldh2*^{-/-} *Fancd2*^{-/-} llegan a nacer, lo que demuestra que el catabolismo materno de los aldehídos es suficiente para rescatar la letalidad embrionaria. Las HSC de los ratones *Aldh2*^{-/-} *Fancd2*^{-/-} sufren un incremento del daño espontaneo al DNA y los ratones desarrollan leucemia aguda y anemia aplásica a lo largo de la vida. Los ratones *Aldh2*^{-/-} *Fancd2*^{-/-} tienen una reducción de 30 veces en el acervo de las HSC respecto a ratones normales y una reducción de 10 veces cuando se les compara con ratones KO solo para *Fancd2*. Además, las HSC de los ratones *Aldh2*^{-/-} *Fancd2*^{-/-} tienen una capacidad de reconstituir la médula ósea de ratones tratados con una dosis letal de radiaciones ionizantes 600 veces inferior con respecto a las HSC de ratones normales (Garaycochea et al., 2012). Cuando experimentos similares se llevan a cabo eliminando *Fanca* (mutado en la mayoría de los pacientes FA) junto con *Aldh2*, el fenotipo resultante es aún peor. Si las madres carecen de ALDH2, los embriones *Aldh2*^{-/-} *Fanca*^{-/-} mueren y son reabsorbidos alrededor de las fases de desarrollo E10-E12.5. Si se implantan en madres *Aldh2*^{+/-}, los embriones doble KO se pueden desarrollar después de la etapa E12.5 pero presentan malformaciones graves, que afectan especialmente a las extremidades y los ojos, y mueren antes de nacer. Solo cuando los embriones *Aldh2*^{-/-} *Fanca*^{-/-} se transfieren a madres *Aldh2*^{+/+} dan lugar a una descendencia viable pero los recién nacidos tienen un bajo número de HSC, lo que indica que el ALDH2 fetal es esencial para una hematopoyesis adecuada (Oberbeck et al., 2014).

Entre el 30 y el 50% de la población asiática de origen China o Japonesa es portadora de un alelo del gen *ALDH2* que presenta el polimorfismo rs671 (c.1510G>A; p.Glu504Lys). La *ALDH2* p.Glu504Lys se comporta como un dominante negativo (Crabb et al., 1989) y en los pacientes FA japoneses que expresan la *ALDH2* p.Glu504Lys el fenotipo es más severo y córrrela con la aparición más temprana del BMF y con el aumento de las malformaciones relacionadas con FA (Hira et al., 2013). Estos datos proporcionan una fuerte evidencia que los resultados de los estudios en ratones pueden tener una gran relevancia también para los seres humanos. Los individuos que expresan la *ALDH2* p.Glu504Lys y consuman alcohol tienen un incremento del riesgo de desarrollar cáncer de esófago, macrocitosis y anemia macrocítica. Además la flora bacteriana presente en las cavidades orales produce gran cantidad de acetaldehído (Yokoyama et al., 2014). Por lo tanto tratamientos dirigidos a eliminar los aldehídos del organismo podrían ser una estrategia válida para prevenir o retrasar tanto el BMF como la aparición de HNSCC en los pacientes FA (Parmar and D'Andrea, 2012, Chen et al., 2014).

II.2. Clonación de los genes FA

Un importante paso previo a la clonación de los genes FA fue la determinación de los grupos de complementación y para eso se aprovechó la característica fundamental de las células FA de ser extremadamente sensibles a los ICL (Figura 2) (Joenje and Patel, 2001). Antiguamente los grupos de complementación de los pacientes FA se averiguaban fusionando líneas celulares derivadas de distintos pacientes y produciendo híbridos celulares con el doble del patrimonio genético (4N): si el híbrido 4N seguía siendo sensible a los ICL significaba que las células derivaban de pacientes que pertenecían al mismo grupo de complementación. Si el híbrido se volvía resistente a los ICL significaba que los dos pacientes tenían defectos en genes diferentes y que las dos células, al fusionarse, complementaban cada una el defecto genético de la otra. A día de hoy, el subtipaje genético se realiza mediante el análisis de la reversión del fenotipo celular una vez introducido en las células del paciente una copia salvaje de cada uno de los genes FA utilizando vectores retro o lentivirales (Hananberg et al., 2002, Casado et al., 2007).

Los experimentos iniciales de determinación de los grupos de complementación demostraron que la FA podía ser producida por alteraciones en unos cuantos genes diferentes relacionados con los grupos de complementación FA-A, B, C, D, E, F, G, I y J (Strathdee et al., 1992a, Joenje et al., 1995, Joenje et al., 1997, Levitus et al., 2004).

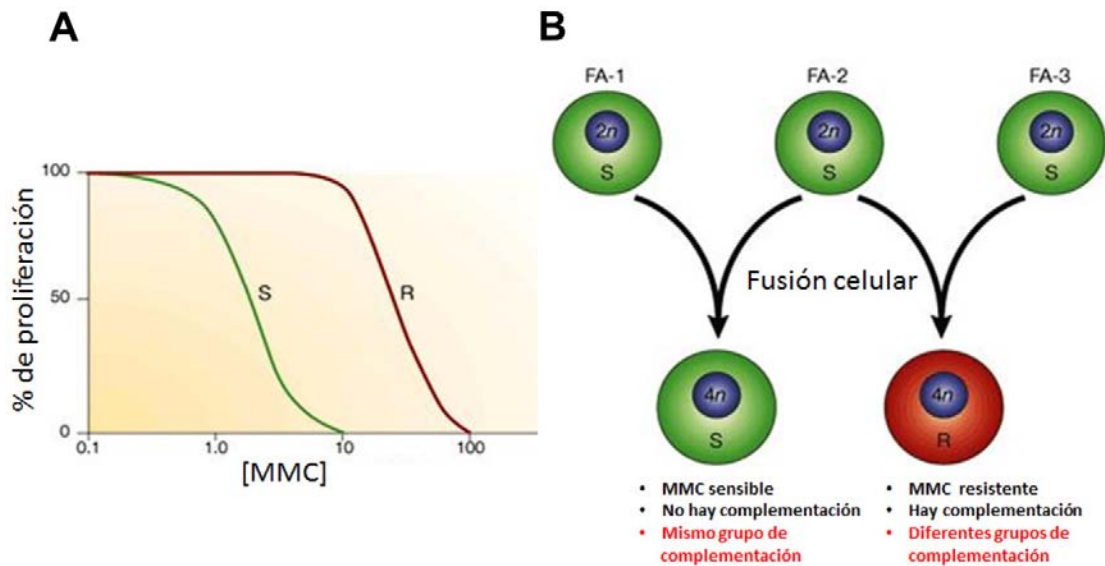


Figura 2: (A) linfoblastos inmortalizados derivados de los pacientes de FA son típicamente desde 10 a 100 veces más sensibles (S) a la acción anti-proliferación de la MMC y de otros agentes que producen ICL, respecto a células no FA o complementadas (R). **(B)** Un híbrido celular, generado por la fusión de linfoblastos de pacientes FA no relacionados, podría continuar a mostrar el fenotipo sensible (S) o no (R). En este ejemplo, FA-1 y FA-2 están en el mismo grupo de complementación, que es diferente al grupo al que pertenece FA-3. Adaptada de (Joenje and Patel, 2001)

Para dilucidar el defecto molecular, las estrategias de clonación utilizadas al principio se basaban en la posibilidad de complementar el fenotipo de sensibilidad a los ICL a través de la introducción de librerías de cDNA en las células FA y seleccionar con MMC las células después de la transfección (Figura 3). A esta técnica se llamó “clonación por complementación. Las células que se habían vuelto resistentes se analizaban después para identificar el cDNA que había complementado el fenotipo de sensibilidad a los ICL. Con esta técnica se identificaron los genes *FANCC*, *FANCA*, *FANCE*, *FANCF* y *FANCG* (Strathdee et al., 1992b, Lo Ten Foe et al., 1996, de Winter et al., 1998, de Winter et al., 2000, de Winter and Joenje, 2009). Lamentablemente, las proteínas codificadas por estos genes eran proteínas hasta al momento desconocidas y no contenían dominios que pudiesen dar pistas sobre de las

funciones que realizaban en las células. Además, las proteínas identificadas aparecieron ser específicas de los vertebrados, ya que no se encontraban ortólogos fuera de este subfilo de los eucariotas. Solamente se pudo averiguar que las diferentes proteínas se organizaban en un complejo proteico sin función conocida (Garcia-Higuera et al., 1999, Garcia-Higuera et al., 2000).

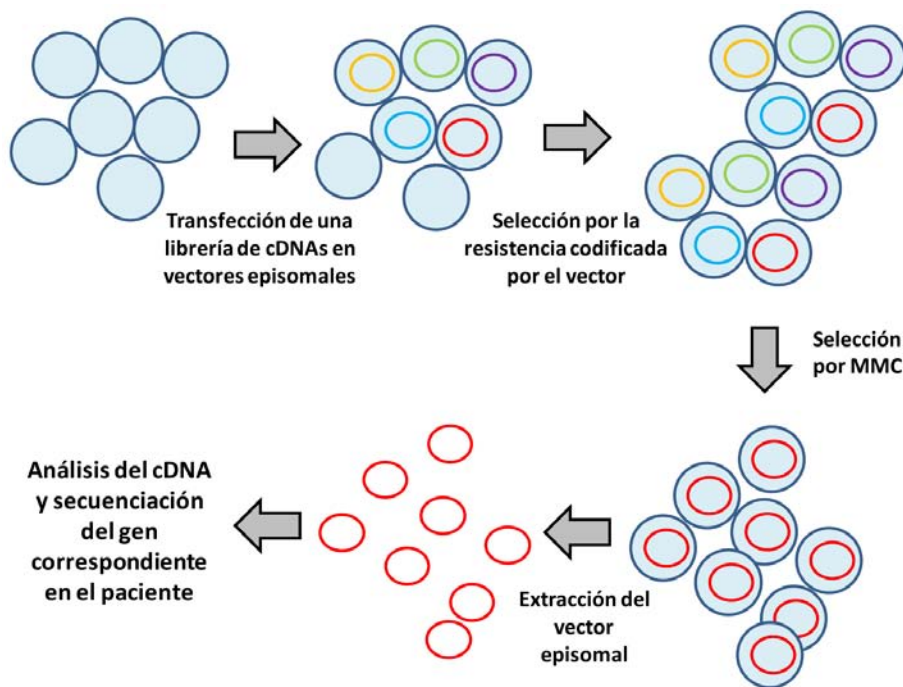


Figura 3: Esquema de la estrategia de clonación por complementación. Linfoblastos FA sensibles a la MMC son transfectados con una librería episomal de cDNA que contiene un marcador de selección. Las células son cultivadas en presencia del agente de selección y las células seleccionadas, a continuación, son cultivadas en medio con una dosis de MMC que mata a las células de FA no corregidas y selecciona las células en las que el defecto se complementa por un cDNA presente en la librería. Posteriormente, el vector episomal se extrae de las células y se analiza el cDNA. Finalmente, el gen correspondiente se analiza en la línea celular del paciente para verificar la presencia de mutaciones patogénicas.

La técnica del “clonaje posicional”, fue utilizada con éxito para descubrir el gen defectuoso en el grupo de complementación D; utilizando la técnica de Transferencia de Cromosomas Mediada por Microcélulas (MMCT) ya se había demostrado que el fenotipo de sensibilidad a la MMC de la línea celular Pd20 (perteneciente al grupo de complementación D) se corregía introduciendo el brazo p del cromosoma 3 en estas células (Whitney et al., 1995). Estudiando por FISH las deleciones en esta región en varios híbridos, se consiguió reducir la región donde estaba el gen *FANCD* (Hejna et al., 2000) y el análisis por

secuenciación de los genes incluidos en la región, identificó el gen afectado (Timmers et al., 2001). Sorprendentemente en dos de los pacientes asignados al grupo de complementación D, este gen no se encontraba mutado y por lo tanto, el grupo de complementación D tuvo que dividirse en dos grupos: FA-D1 y FA-D2. El nuevo gen se encontraba mutado solo en los pacientes pertenecientes al grupo FA-D2 y por lo tanto fue llamado *FANCD2*. El descubrimiento de *FANCD2* fue de capital importancia porque aportó por primera vez un claro indicio que en las células de la FA había un defecto en la reparación del DNA: después de daño al DNA la proteína FANCD2 se localizaba en *foci* nucleares con el producto del gen de susceptibilidad de cáncer de mama *BRCA1/FANCS* y la proteína clave para la Reparación del DNA por Recombinación Homóloga (HRR) RAD51/FANCR. Esta localización se producía durante la fase S del ciclo celular y requería la monoubiquitinación de FANCD2 en la Lisina 561 (K561). La monoubiquitinación de FANCD2 era estrictamente dependiente de la presencia de proteínas FANCA, C, F y G funcionales. Estas observaciones sugerían que la FA pudiese ser causada por un defecto de la HRR durante la fase S del ciclo celular (García-Higuera et al., 2001, Taniguchi et al., 2002a) y la estrecha relación entre FANCD2 y BRCA1/FANCS llevó a investigar los genes *BRCA1/FANCS* y *BRCA2/FANCD1* en los pacientes FA-D que carecían de mutaciones en *FANCD2*. Con este enfoque por genes candidatos, se descubrió que los pacientes del grupo de complementación FA-D1 tenían mutaciones bialélicas en *BRCA2/FANCD1* y expresaban proteínas BRCA2/FANCD1 truncadas (Howlett et al., 2002). Este descubrimiento conectó definitivamente la FA a la HRR y a los genes de susceptibilidad al cáncer de mama y, para enfatizar esta conexión, la ruta FA fue rebautizada como ruta FA/BRCA. Seguía pero siendo un misterio lo que el complejo nuclear de FANCA,-C,-E,-F y G en realidad estaba haciendo y cómo estaba involucrado en el monoubiquitinación de FANCD2.

Usando un anticuerpo contra FANCA se purificó por inmunoprecipitación (IP) de extractos nucleares de células HeLa un complejo formado de muchas proteínas que luego se analizó por espectrometría de masas (Figura 4). Se descubrieron así varias proteínas nuevas que eran también parte del complejo formado por FANCA,-C,-E,-F y G, que se llamaron Proteínas Asociadas a la Anemia de Fanconi (FAAPs) seguido por su peso molecular en KDa. Dos de ellas, la FAAP43 y la FAAP95, resultaron ser genuinas proteínas FA porque los genes que las codificaban estaban mutados en pacientes de los grupos de complementación L

y B. La FAAP43 y la FAAP95 pasaron entonces a llamarse FANCL y FANCB respectivamente.

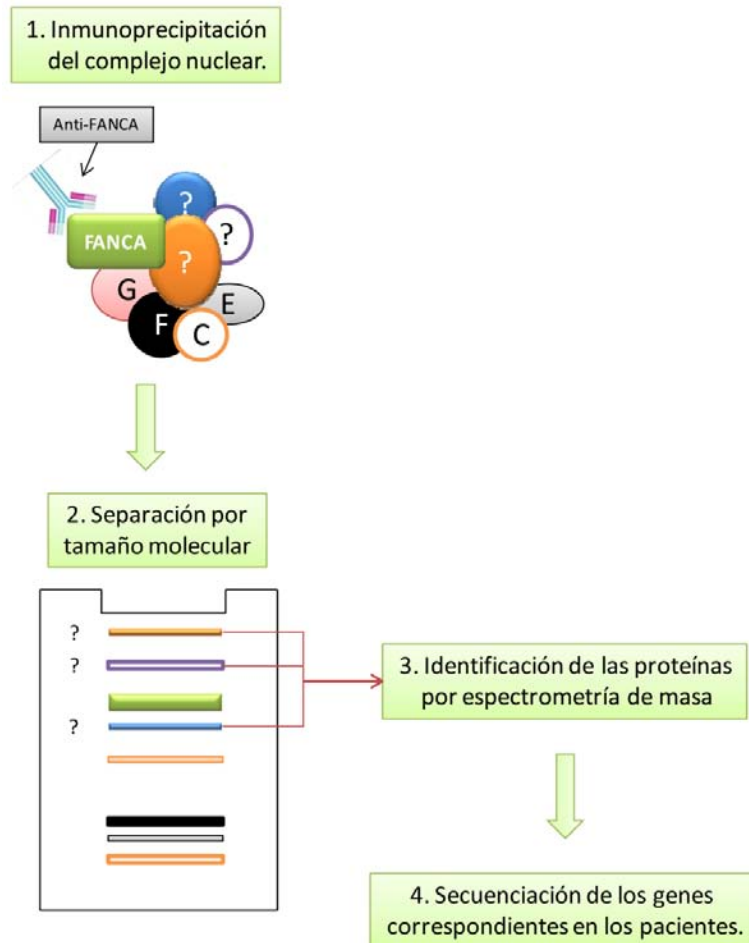


Figura 4: Identificación de genes FA por estudios de co-IP. Un anticuerpo contra FANCA se añade a un extracto nuclear de células HeLa para precipitar FANCA y las proteínas que se unen a ella. Se separan las proteínas precipitadas por SDS-PAGE y se analizan por espectrometría de masas. Posteriormente, se analiza a los genes que codifican las proteínas nuevas en los pacientes FA de los cuales se desconoce el gen mutado.

Muy importante fue el descubrimiento que FANCL era una ubiquitin-E3 ligasa porque por fin se podía conectar directamente el complejo nuclear FA con la monoubiquitinación de FANCD2. La característica principal de *FANCB* era de ser el único gen FA que tenía una transmisión ligada al sexo (Meetei et al., 2003, Meetei et al., 2004) al estar localizado en el cromosoma X. Otra FAAP identificada con este sistema fue la

FAAP250, el ortólogo humano de la nucleasa Hef de las arqueobacterias, pero, en contraste con Hef, FAAP250 no tenía el dominio nucleasa funcional (Mosedale et al., 2005). En un primer momento se encontró un paciente FA que parecía tener mutaciones patogénicas en el gen y por lo tanto se convirtió en *FANCM*, luego se vio que dicho paciente tenía también mutaciones en *FANCA* pero el nombre de *FANCM* se mantuvo aunque no se hayan encontrado todavía pacientes FA con solo mutaciones en *FANCM* (Meetei et al., 2005, Singh et al., 2009).

BRIP1/FANCI fue identificado por MMCT y por análisis de ligamiento genético. El análisis de ligamiento con marcadores polimórficos a lo largo de todo el genoma en un paciente FA-J perteneciente a una familia consanguínea, reveló una extensa región en homocigosis en el cromosoma 17 y experimentos de MMCT confirmaron que el cromosoma 17 corregía el fenotipo FA de las células del paciente. Dentro de la región homocigótica el gen candidato más convincente era *BRIP1/FANCI* que codificaba por una helicasa que interactuaba con BRCA1/FANCS y cuya ausencia confería a las células un fenotipo FA. La secuenciación del DNA reveló mutaciones patogénicas en *BRIP1/FANCI* en todos los pacientes clasificados como FA-J reforzando aún más el modelo de la ruta FA/BRCA para la reparación del DNA (Levitus et al., 2005, Levrán et al., 2005a, Litman et al., 2005).

De forma parecida a la purificación del complejo nuclear, en un experimento a gran escala de co-IP se investigaron las proteínas que se unían a BRCA2/FANCD1 y se identificó así a PALB2, una nueva proteína que resultó ser esencial para la asociación de BRCA2/FANCD1 con la cromatina y para sus funciones en la HRR (Xia et al., 2006). Las células deficientes para PALB2 tenían fenotipo FA y, de hecho, varios pacientes de FA aún no clasificados tenían mutaciones patogénicas en *PALB2*. Por lo tanto se rebautizó *PALB2* como *FANCD2* (Xia et al., 2007, Reid et al., 2007).

FANCI fue encontrada en un cribado a gran escala de sustratos de ATM/ATR utilizando anticuerpos fosfo-específicos y espectrometría de masas (Smogorzewska et al., 2007), al mismo tiempo que se identificaba también en un análisis *in silico* para posibles sustratos de monoubiquitinación del complejo nuclear FA (Sims et al., 2007) y por análisis de ligamiento en cuatro familias FA-I (Dorsman et al., 2007). Se vio que la proteína FANCI

se monoubiquitinaba de forma parecida a FANCD2 y que las dos proteínas formaban un dímero estable que se bautizó como “complejo ID” (Smogorzewska et al., 2007).

En dos hermanos de origen paquistaní, hijos de padres consanguíneos, se había identificado un síndrome caracterizado por múltiples anomalías congénitas que eran características de FA. Las células derivadas de los pacientes tenían un clásico fenotipo FA pero los dos hermanos, notablemente, no presentaban anemia. Utilizando un array de polimorfismos del DNA de un único nucleótido (*Single Nucleotide Polymorphisms*; SNP) se identificó en el cromosoma 17 de los dos hermanos una región en homocigosis (17q21–q24) que contenía tres genes involucrados en la reparación del DNA: *EME1*, *RAD51C* y *BRIP1/FANCI*. El análisis por secuenciación del DNA demostró que una mutación de cambio de sentido en homocigosis en el gen *RAD51C* es responsable del síndrome que padecían los dos hermanos (Vaz et al., 2010). El gen *RAD51C* se rebautizó *FANCO* aunque hay que remarcar que no hay un consenso en la comunidad científica sobre si *RAD51C* es un verdadero gen FA debido a la ausencia de anemia en los pacientes estudiados (Garner and Smogorzewska, 2011).

SLX4 se describió por primera vez como una proteína que proporcionaba un soporte para el anclaje de tres nucleasas diferentes (SLX1, ERCC4/XPF-ERCC1 y MUS81-EME1) coordinando de esta manera la resolución de las uniones de Holliday y la reparación de los ICL en la fase S del ciclo celular (Fekairi et al., 2009, Holloway et al., 2011, Munoz et al., 2009, Andersen et al., 2009). El hecho que las células deficientes para SLX4 tuvieran un fenotipo de sensibilidad a los ICL lo convirtió en un candidato para ser un gen involucrado en la FA. El análisis del DNA confirmó que mutaciones en el gen *SLX4* eran causa de FA en los pacientes pertenecientes al grupo de complementación FA-P y por lo tanto a *SLX4* se le identificó como *FANCP* (Stoepker et al., 2011, Kim et al., 2011). Además, se demostró en ratón que la interacción de Slx4 con Xpf/Ercc1 era indispensable para corregir el fenotipo de sensibilidad a los ICL (Crossan et al., 2011), un dato cuya importancia luego se confirmó, en el contexto de esta tesis doctoral, cuando se identificó por secuenciación del exoma a *ERCC4* como el gen *FANCO* (Bogliolo et al., 2013).

Desde el descubrimiento de *BRCA2/FANCD1* como un gen FA, *BRCA1/FANCS* siempre estuvo en el punto de mira como candidato para ser también un gen FA. La falta de

individuos con mutaciones bialélicas en *BRCA1/FANCS*, no obstante en algunas poblaciones haya una prevalencia elevada de portadores (hasta 1 de cada 40 en la población de los judíos Ashkenazi (Petrucci et al., 2010)), siempre se había creído ser indicativa que las mutaciones bialélicas en *BRCA1/FANCS* resultaban letales a nivel embrionario. En 2013 se identificó un primer paciente con mutaciones bialélicas en *BRCA1/FANCS* que tenía rasgos fenotípicos de FA pero, al faltar una línea celular derivada del paciente y los estudios del fenotipo celular y de complementación, *BRCA1/FANCS* no se había incorporado aun de forma oficial a la lista de los genes FA (Domchek et al., 2013). Pero en 2014, se publicó un segundo paciente (Sawyer et al., 2014) y esta vez, al haber una línea celular disponible, se pudieron completar todos los estudios formales confirmando que mutaciones bialélicas en *BRCA1/FANCS* producían un síndrome parecido a FA, pero, como en el caso de *RAD51C/FANCO*, ninguno de los dos pacientes tenía BMF.

La “llave de paso” de la ruta FA/BRCA es la monoubiquitinación del complejo ID por parte del FANCore a través de la función catalítica de FANCL que actúa como una E3-ubiquitina ligasa (Meetei et al., 2003, Alpi et al., 2008, Longerich et al., 2009). La UBE2T es el enzima E2 de conjugación de ubiquitina esencial para la monoubiquitinación de FANCD2 (Machida et al., 2006) y FANCI (Longerich et al., 2009) y tres grupos distintos muy recientemente identificaron mutaciones bialélicas en *UBE2T* como causa de FA (Virts et al., 2015, Rickman et al., 2015, Hira et al., 2015). En total se han identificado 4 pacientes y en dos de ellos se observó BMF (Hira et al., 2015). Los otros dos pacientes resultaron ser mosaicos por la reversión de una de las mutaciones a nivel del tejido hematopoyético explicando así la ausencia de un claro fenotipo hematológico en estos dos últimos pacientes (Rickman et al., 2015, Virts et al., 2015). *UBE2T* así se rebautizó con el alias de *FANCT*

El último gen FA publicado hasta al momento es *RAD51/FANCR*. En dos pacientes distintos se han visto como mutaciones monoalélicas en *RAD51* con efecto dominante negativo causaban un síndrome parecido a FA, pero, como en el caso de *RAD51C/FANCO* y de *BRCA1/FANCS*, ninguno de los dos pacientes presenta un defecto a nivel de la hematopoyesis (Wang et al., 2015).

II.3. Genes FA, FA-like y asociados a FA

Diecinueve genes se han visto hasta ahora estar involucrados en FA pero, siguiendo un criterio más estricto que considere que un gen para ser llamado FA tiene que encontrarse mutado como mínimo en dos pacientes y que por lo menos uno de ellos tenga BMF, solo 15 de estos genes cumplen para ser considerados verdaderos genes FA (*FANCA, B, C, DI, D2, E, F, G, I, J, L, N, P, Q y T*) (Tabla II). *FANCO, S y R* se deberían considerar genes “FA-like” porque se encuentran mutados en un síndrome de fragilidad cromosómica y malformaciones típicas de FA pero sin BMF u otras alteraciones a nivel de la sangre (Bogliolo and Surralles, 2015).

Tabla II: Genes con mutaciones bialélicas en FA: En azul se indican los genes FA y en verde los genes FA-like.

Gen FA	Nombre alternativo	Localización cromosómica	Referencias
FANCA	-	16q24.3	(Lo Ten Foe et al., 1996)
FANCB	-	Xp22.2	(Meetei et al., 2004)
FANCC	-	9q22.32	(Strathdee et al., 1992b)
FANCD1	BRCA2	13q13.1	(Howlett et al., 2002)
FANCD2	-	3p25.3	(Garcia-Higuera et al., 2001)
FANCE	-	6p21.31	(de Winter et al., 2000a)
FANCF	-	11p14.3	(de Winter et al., 2000b)
FANCG	XRCC9	9p13.3	(de Winter et al., 1998)
FANCI	-	15q26.1	(Smogorzewska et al., 2007)
FANCI	BRIP1	17q23.2	(Levitus et al., 2005)
FANCL	PHF9	2p16.1	(Meetei et al., 2003)
FANCN	PALB2	16p12.2	(Xia et al., 2007)
FANCO	RAD51C	17q22	(Vaz et al., 2010)
FANCP	SLX4	16p13.3	(Stoepker et al., 2011, Kim et al., 2011)
FANCQ	ERCC4	16p13.12	(Bogliolo et al., 2013)
FANCT	UBE2T	1q32.1	(Virts et al., 2015, Rickman et al., 2015, Hira et al., 2015)
FANCR	RAD51	15q15.1	(Wang et al., 2015)
FANCS	BRCA1	17q21	(Sawyer et al., 2014)

FANCM no se debería considerar un gen FA ni un gen FA-like: el único paciente de FA conocido que presentaba mutaciones bialélicas en *FANCM* también tenía mutaciones bialélicas y patogénicas en *FANCA* (Singh et al., 2009), el cDNA de *FANCM* no revertía el defecto genético de la células del paciente (Meetei et al., 2005) y el hermano, que también presentaba FA, fue subtipado FA-A (Singh et al., 2009). Además variantes de pérdida de función de *FANCM* son más frecuentes de lo esperado en la población general y se han

identificado dos individuos Finlandeses sanos y con perfil hematológico normal que son homocigotos para mutaciones dañinas en *FANCM* (Lim et al., 2014). Por lo tanto sería mejor considerar a *FANCM* como una FAAP que como la FAAP100, la FAAP24, la FAAP20, MHF1 (FAAP16) y MHF2 (FAAP10) participa en la ruta de FA/BRCA de reparación del DNA aunque no esté involucrada en la etiología de la enfermedad (Bogliolo and Surralles, 2015).

II.4. Biología molecular de la FA

Los ICL son lesiones al DNA de extrema gravedad porqué unen de forma covalente a las dos cadenas del DNA y, por lo tanto, impiden procesos fundamentales como la replicación y la transcripción. Además, afectando a ambas las cadenas hacen que no haya una hebra de DNA libre de daño para ser utilizada como molde en un proceso de reparación libre de errores. La ruta FA/BRCA coordina las reacciones necesarias para la eliminación del ICL y el restablecimiento de la integridad del genoma (Figura 5). Siete proteínas FA (*FANCA*, *FANCB*, *FANCC*, *FANCE*, *FANCF*, *FANCG*, *FANCL*) y 6 FAAPs (*FANCM/FAAP250*, *FAAP100*, *FAAP24*, *FAAP20*, *FAAP16* y *FAAP10*) forman un complejo conocido como el complejo nuclear Fanconi (FANCore Complex o FANCore) que se activa en respuesta a la horquillas de replicación bloqueadas y monoubiquitina a ambos componentes del complejo ID (Garcia-Higuera et al., 2001, Smogorzewska et al., 2007, Kaddar and Carreau, 2012, Wang and Smogorzewska, 2015). El complejo ID activado se localiza entonces en el sitio del daño (Garcia-Higuera et al., 2001, Bogliolo et al., 2007, Smogorzewska et al., 2007) y, a lo largo de un proceso todavía por aclarar, promueve la actividad nucleasa para cortar el DNA a nivel del ICL y la sucesiva síntesis a través de la lesión del DNA (*Translesion synthesis*; TLS) que queda en el DNA molde (Raschle et al., 2008, Knipscheer et al., 2009). Sucesivamente la lesión acaba de ser reparada por HRR y se restablece así la correcta información en el DNA (Figura 5). Además de activarse por daños exógenos al DNA, la ruta FA/BRCA se activa también durante la normal fase S del ciclo celular, lo que sugiere que podría participar también en la reparación de los errores que se producen a lo largo de la fase S y que bloquean la progresión de las horquillas de replicación (Garcia-Higuera et al., 2001, Taniguchi et al., 2002a).

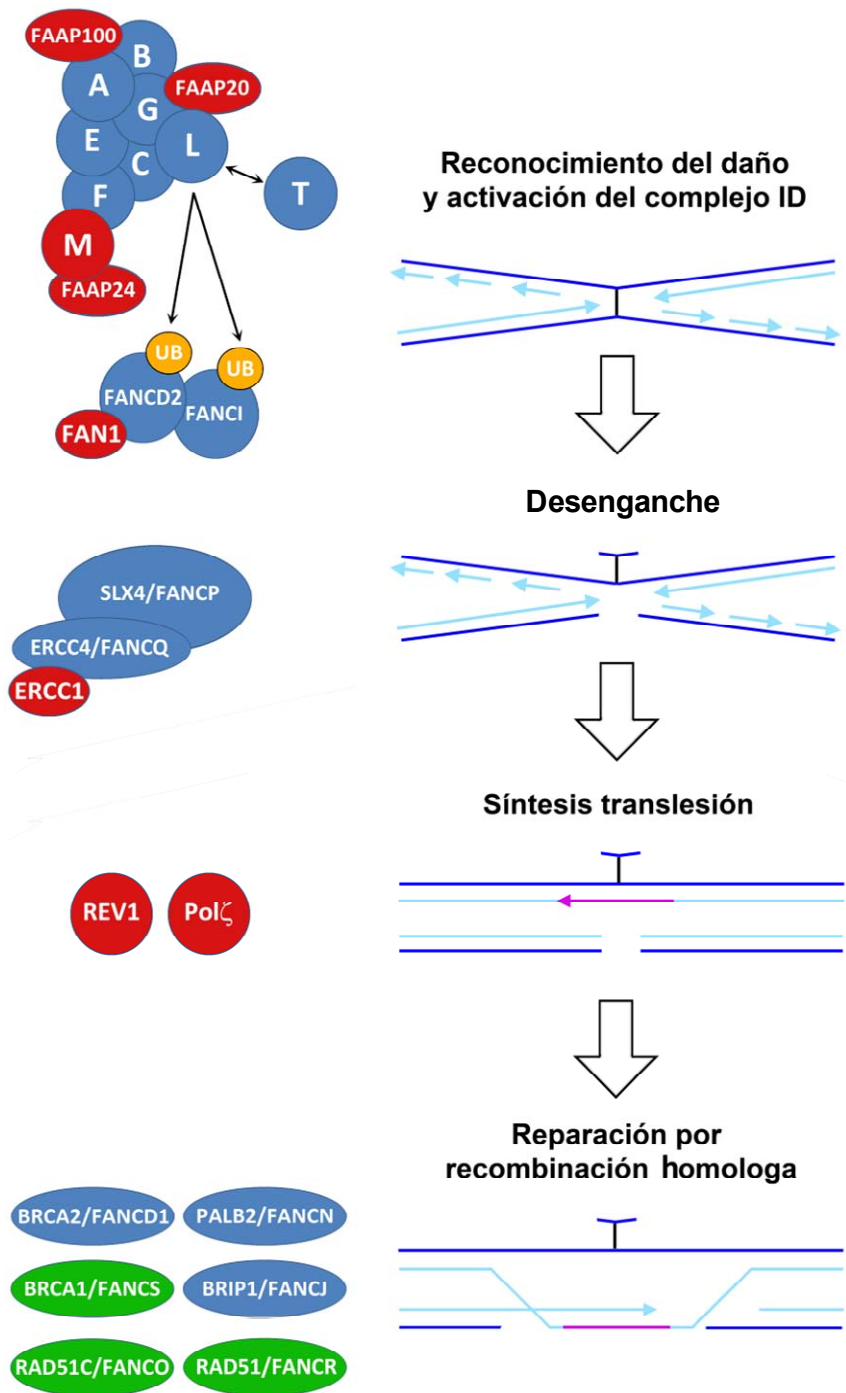


Figura 5: La vía FA/BRCA de reparación de los ICL: El bloqueo de las horquillas de replicación por parte de un ICL activa el complejo nuclear FA que monoubiquitina al complejo ID. El complejo ID activado se traslada a la cromatina y promueve la actividad nucleasa del complejo SLX4-XPF/ERCC1 que cortan en proximidad de la lesión desenganchando así las dos cadenas del DNA y permitiendo la síntesis translesión. En el paso final de este proceso, la integridad del genoma es restaurada por recombinación homóloga gracias a la acción de las proteínas FA que actúan en esta última etapa. Las proteínas codificadas por los genes FA se muestran en azul, mientras que los productos de los genes FA-like se muestran en verde. Las proteínas asociadas a FA cuyos genes no están mutados en ningún paciente FA o FA-like se muestran en rojo (modificada de (Bogliolo and Surrallés, 2015)).

Tanto la monoubiquitinación de FANCD2 como la de FANCI están ausentes en células que llevan mutaciones en cualquiera de los genes cuyas proteínas constituyen el FANCore o UBE2T, con la excepción de *FANCM*. Aunque *FANCM* haya sido reportado como un componente esencial del FANCore, es prescindible para la monoubiquitinación del complejo ID (Meetei et al., 2005, Singh et al., 2009). La monoubiquitinación de FANCD2 en la K561 es tan esencial para las funciones de la ruta FA que las células que expresan el mutante K561R que no puede ser monoubiquitinado, son tan sensibles a los ICL como las células nulas para FANCD2 (Garcia-Higuera et al., 2001). La ubiquitinación es una modificación postraduccional en la cual el último aminoácido (Glicina 76) de la ubiquitina se une a un residuo de Lisina de una proteína sustrato. Se puede añadir una sola molécula de ubiquitina o una cadena de moléculas (Dikic and Robertson, 2012) y estas modificaciones son un evento de señalización versátil y con una amplia gama de posibles consecuencias. La poli-ubiquitinación generalmente destina una proteína para la degradación por el proteosoma (Glickman and Ciechanover, 2002), pero la monoubiquitinación cumple un rol no proteolítico y, en los últimos años, se han acumulado evidencias que varias proteínas involucradas en la reparación del DNA se ubiquitinan en respuesta al daño al DNA. Típicamente, estas proteínas se localizan cerca del sitio de la lesión del DNA en *foci* nucleares como por ejemplo PCNA y la histona H2AX, dos proteínas clave cuya ubiquitinación sirve para reclutar los factores necesarios para la respuesta celular al daño al DNA (Hofmann, 2009, Panier and Durocher, 2009, Ulrich, 2009b, Ulrich, 2009a). La ubiquitinación requiere tres tipos de enzimas: enzimas que activan la ubiquitina (E1), enzimas que acoplan la ubiquitina al sustrato (E2) y ligasas de ubiquitina (E3). FANCL representa la ligasa E3 necesaria para la conjugación de la ubiquitina a la K561 de la FANCD2 y a la K523 de FANCI en humanos (Cole et al., 2010). Experimentos con doble híbrido de levadura han identificado dos proteínas E2 capaces de interactuar con FANCL *in vitro*, UBE2T y UBE2W (Alpi et al., 2008) pero solo la UBE2T es la enzima E2 para la ubiquitinación de FANCD2 dado que la UBE2T es necesaria también para la resistencia celular a los ICL *in vivo* (Machida et al., 2006, Hira et al., 2015, Rickman et al., 2015, Virts et al., 2015). FANCD2 y FANCI son de momento los únicos sustratos conocidos del FANCore aunque no se pueda descartar que haya otras proteínas ubiquitinadas por esta compleja ligasa E3. La contribución de la monoubiquitinación de FANCI a la resistencia

celular a los ICL parece ser menos importante que la de FANCD2 visto que el mutante de monoubiquitinación de FANCI es capaz de revertir el fenotipo de sensibilidad a los ICL de las células deficientes en FANCI como si fuera la proteína normal (Alpi et al., 2008, Ishiai et al., 2008). Igualmente importante es el evento de apagar esta señal una vez concluida la reparación del DNA. La falta de de-ubiquitinación del complejo ID también resulta en sensibilidad a los ICL (Oestergaard et al., 2007, Kim et al., 2009). Esto implica que el ciclo del complejo ID entre las formas ubiquitinada y no ubiquitinada es esencial para la actividad de la ruta. USP1 es la deubiquitinasa que elimina la monoubiquitinación de FANCD2 (Nijman et al., 2005). Los niveles de USP1 están estrictamente regulados a nivel de la transcripción y hay un pico durante la fase S, presumiblemente cuando su actividad es más necesaria (Nijman et al., 2005) mientras que después de daño al DNA, la transcripción de *USP1* se apaga con un rápido descenso de los niveles de la proteína (Cohn et al., 2007). Curiosamente, USP1 es también la enzima que de-ubiquitina a PCNA, sugiriendo que los procesos de deubiquitinación de PCNA y del complejo ID podrían tener una regulación común (Huang et al., 2006).

Después del daño al DNA, las proteínas de la ruta FA/BRCA también están reguladas por las quinasas ATM y ATR: ATM se activa en respuesta a roturas de doble cadena (*double strand breaks*; DSBs) mientras que ATR lo hace en respuesta a señales de estrés replicativo (Smith et al., 2010). Varios componentes de la vía de la FA son fosforilados por ATM y ATR incluyendo FANCA, FANCG, FANCM, FANCD1 y ambos componentes del complejo ID (Ishiai et al., 2008, Smogorzewska et al., 2007, Taniguchi et al., 2002b, Meetei et al., 2005, Ho et al., 2006, Mi et al., 2004, Qiao et al., 2004, Matsuoka et al., 2007). La importancia de muchos de los eventos de fosforilación no está enteramente aclarada, pero algunos de ellos son importantes para la resistencia a los ICL. En particular, la fosforilación de FANCI en sitios SQ/TQ cerca de su sitio de monoubiquitinación es necesaria para la monoubiquitinación y la localización del complejo ID a los *foci* nucleares después de daño al DNA. Basándose en estos resultados, se ha propuesto que la fosforilación de FANCI funcione como un interruptor molecular de la vía FA/BRCA (Ishiai et al., 2008).

La etapa de incisión del ICL para “desenganchar” (*unhooking*) el DNA es fundamental, porque sin ella la reparación del ICL resultaría totalmente imposible (Figura 5). En este paso la proteína FANCP (SLX4) juega un papel esencial. La SLX4 es una gran

proteína, de alrededor 200 KDa, que funciona como un soporte para coordinar las actividades de por los menos tres endonucleasas distintas (XPF/ERCC1, MUS81/EME1 y SLX1) involucradas a varios niveles en la resistencia a los ICL y en la resolución de las uniones de Holliday (Cybulski and Howlett, 2011, Munoz et al., 2009). Las células derivadas de los pacientes FA-P monoubiquitan normalmente a FANCD2 después de daño al DNA sugiriendo que SLX4 actúa después de la activación del complejo ID. La parte N-terminal de SLX4 tiene dos dominios de interacción con la ubiquitina (*ubiquitin-binding zinc finger domain*; UBZ)) y las mutaciones en los pacientes sugieren un papel importante de estos dominios para las funciones de SLX4 (Kim et al., 2011, Stoepker et al., 2011). Esto podría indicar que FANCP podría localizar a los sitios de daño vía interacciones con proteínas monoubiquitinadas pero los sustratos de estas interacciones se desconocen en este momento. Estudios en modelos de ratones KO para *Slx4* demostraron además que la interacción entre *Slx4* y *Xpf/Ercc1* es esencial para revertir el fenotipo de sensibilidad a los ICL de las células derivadas de los ratones *Slx4*^{-/-}, indicando a XPF/ERCC1 como la endonucleasa más importante en la reparación de los ICL (Crossan et al., 2011). Los descubrimientos sucesivos que mutaciones bialélicas en *ERCC4* causan FA (Bogliolo et al., 2013) y que XPF/ERCC1 es la única endonucleasa realmente necesaria para la incisión nucleolítica cerca de los ICL que bloquean las horquillas de replicación del DNA (Klein Douwel et al., 2014) corroboran definitivamente que XPF/ERCC1 es la endonucleasa principal de la ruta FA/BRCA para la reparación del DNA. Una vez abierto el DNA por parte del complejo XPF/ERCC1-SLX4 los modelos más consensuados de reparación de los ICL prevén una fase de TLS por encima de la base donde el enlace entrecruzado ha quedado “colgando” (Figura 5) para que se forme un sustrato con una rotura de doble cadena del DNA (*double strand break*; DSB) que se pueda reparar por HRR. La función especializada de las polimerasas translesión REV1 y Polζ en la reparación de los ICL salió a la luz cuando se vio como las células de pollo DT40 que carecían de una de estas polimerasas eran extremadamente sensibles a los ICL (Niedzwiedz et al., 2004). Además REV1 y Polζ son epistáticas con el gen FANCC en el modelo celular DT40 al no haber prácticamente diferencias entre los mutantes para FANCC y los dobles mutantes FANCC/REV1 o FANCC/polζ en términos de sensibilidad a los ICL (Niedzwiedz et al., 2004, Nojima et al., 2005). Posteriormente, en un sistema de extracto de huevos de *Xenopus laevis*, se demostró que el complejo ID activado promueve la incisión y la TLS en la

reparación de los ICL (Knipscheer et al., 2009). Según este estudio, la monoubiquitinación del complejo ID podría desempeñar un papel en la TLS o por interacción directa con los dominios UBZ presentes en REV1 y pol ζ o a través de una interacción directa entre FANCD2 y PCNA ubiquitinada (Howlett et al., 2009). El modelo propuesto es que REV1 inserte un nucleótido en frente del ICL que ha sido desenganchado, y que Pol ζ a través de su subunidad catalítica REV3 extienda la nueva cadena de DNA (Raschle et al., 2008, Knipscheer et al., 2009). La falta de FANCD2 afecta tanto al paso de incisión previa como a lo de TLS pero no se sabe todavía si ambas actividades son reguladas por el complejo ID o si el complejo ID regula solamente una de ellas y la segunda depende de la iniciación y/o la terminación de la primera. Una vez completado el paso de TLS la reparación se completa por HRR clásica con la intervención de las proteínas FA directamente involucradas en este último paso (Deans and West, 2011).

II.4.1. FAN1

La “nucleasa asociada a FA 1” (FAN1) fue identificada en 2010 e inmediatamente monopolizó toda la atención como el principal candidato para ser la nucleasa que actuaba en la etapa de unhooking de la ruta FA/BRCA (MacKay et al., 2010, Kratz et al., 2010, Smogorzewska et al., 2010, Liu et al., 2010). FAN1 es una nucleasa que reconoce estructuras específicas de DNA y su inhibición induce sensibilidad celular a los ICL. En estos estudios se vio como FAN1 era reclutada al DNA dañado a través de la interacción de su dominio UBZ N-terminal con la FANCD2 monoubiquitinada. El reclutamiento de FAN1 a través de la FANCD2 monoubiquitinada era consistente con la pérdida de las incisiones nucleolíticas cerca del ICL que se observaba *in vitro* al eliminar a FANCD2 (Knipscheer et al., 2009). Pero, sucesivamente, se demostró que la deficiencia de FAN1 no estaba asociada con un fenotipo FA (Trujillo et al., 2012), aunque FAN1 interactúe con FANCD2 y contribuya a la ruta FA/BRCA (Wang et al., 2014, Zhao et al., 2014, Chaudhury et al., 2014). La actividad nucleasa de FAN1 (como la actividad de MUS81-EME1) no se requiere para las incisiones nucleolíticas cerca de un ICL (Klein Douwel et al., 2014) y mutaciones bialélicas que inhiben la función de FAN1 causan una enfermedad (Nefritis intersticial) de insuficiencia renal

(Zhou et al., 2012). Por lo tanto, el papel de FAN1 en la vía FA/BRCA todavía está en discusión.

II.5. Reparación por Recombinación Homóloga y FA

La HRR (Figura 6) repara los DSBs durante la fase S y G2 del ciclo celular. Esta ruta de reparación probablemente evolucionó para solucionar las roturas de DNA que se forman cuando la horquilla de replicación encuentra una lesión en el DNA, se bloquea y sucesivamente colapsa (Figura 6A). La HRR también repara las roturas que se forman después que el DNA acaba de ser replicado (Figura 6B y C) y es vital para salvaguardar la fidelidad de la replicación porque utiliza la cromátida hermana intacta como molde para la síntesis del DNA necesaria para restablecer la integridad de la información genética (Figura 6). En los eucariotas las proteínas BRCA1/FANCS, BRCA2/FANCD1 y RAD51/FANCR juegan un papel fundamental en la HRR y BRCA1/FANCS participa en por los menos cuatro macro complejos (Figura 7) que funcionan a nivel del reconocimiento de la lesión, de la reparación del DNA y de la activación de los puntos de control del ciclo celular (Figura 7) (Livingston, 2009, Greenberg et al., 2006, Roy et al., 2012). En dos de estos macro complejos participan también las otras proteínas FA que actúan en la etapa de HRR de los ICL. Un primer macro-complejo que comprende BRCA1/FANCS, RPA80a y Abraxas localiza al daño en manera dependiente de la fosforilación y de la ubiquitinación de la histona H2AX marcando el comienzo de la HRR (Wang et al., 2007, Kim et al., 2007a). Sucesivamente al sitio de la lesión llegan el complejo MRN, formado por las proteínas MRE11, RAD50 y Nijmegen Breakage Syndrome 1 (NBS1) (Stracker and Petrini, 2011) y otro complejo formado por BRCA1/FANCS y la proteína CtIP. Esta última funciona como un regulador que inhibe la reparación por unión no homóloga de los extremos (*Non Homologous End Joining*; NHEJ) y que favorece en su lugar la HRR (Bunting et al., 2010). El complejo MRN estimulado por la interacción con BRCA1/FANCS-CtIP (Nimonkar et al., 2008, Gravel et al., 2008) actúa como una exonucleasa y elimina una parte de las cadenas del DNA en dirección 5'–3' generando unas extremidades 3' sobresalientes de DNA de cadena sencilla (ssDNA) para la sucesiva invasión y búsqueda de homología en la cromátida hermana (Sartori et al., 2007, Yun and Hiom, 2009, Shibata et al., 2011).

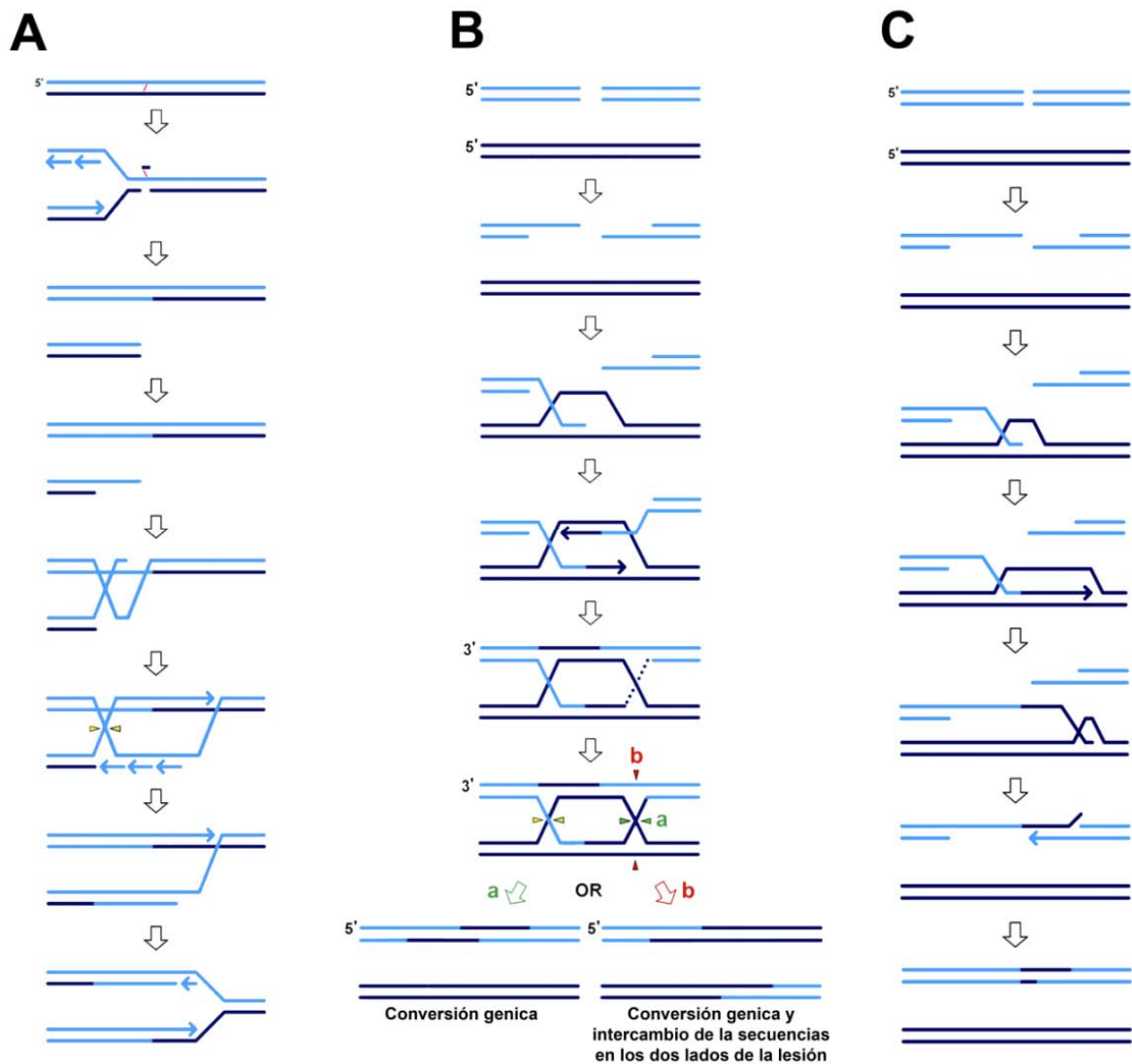


Figura 6. (A) El colapso de una horquilla de replicación produce un DSB con una sola extremidad. La rotura viene modificada para crear en 3' un extremo sobresaliente de ssDNA. El extremo de ssDNA invade la cromátida hermana. Después de haber encontrado la secuencia homóloga se crea el D-loop y por síntesis de DNA se forma una unión de Holliday. La resolución de la unión restablece la horquilla de replicación produciendo un intercambio de cromátidas hermanas (*Sister chromatid Exchange*; SCE). **(B)** Cuando se forma un DSB después que el DNA ha sido replicado, este puede ser reparado también por HRR. El DSB en este caso tiene dos extremidades que tienen que ser modificadas a 3' ssDNA. Una de esta extremidades invade la cromátida hermana como en (A) y se forma el D-loop. La otra extremidad ssDNA se empareja con el D-loop y la síntesis del DNA produce una doble unión de Holliday que se puede resolver en dos maneras distintas: en (a) se preservan las regiones a lado de la lesión y hay conversión génica solo en el punto de la rotura; en (b) se produce conversión génica y un SCE. **(C)** Emparejamiento de la hebra dependiente de síntesis (*synthesis dependent strand annealing*). Esta ruta comienza como en (B) con un DSB con dos extremidades modificadas a ssDNA sobresalientes pero solo una de ellas invade la cromátida hermana y produce el D-Loop. La otra extremidad no se empareja con el D-loop. La síntesis desplaza la hebra "invasora" que finalmente se libera sin la necesidad de cortes y se vuelve a emparejar con la con la otra extremidad a ssDNA. Una nucleasa recorta el DNA que sobra y se restablece la correcta información genética, produciendo conversión génica solo en el punto de la rotura (adaptada de (Helleday et al., 2007)).

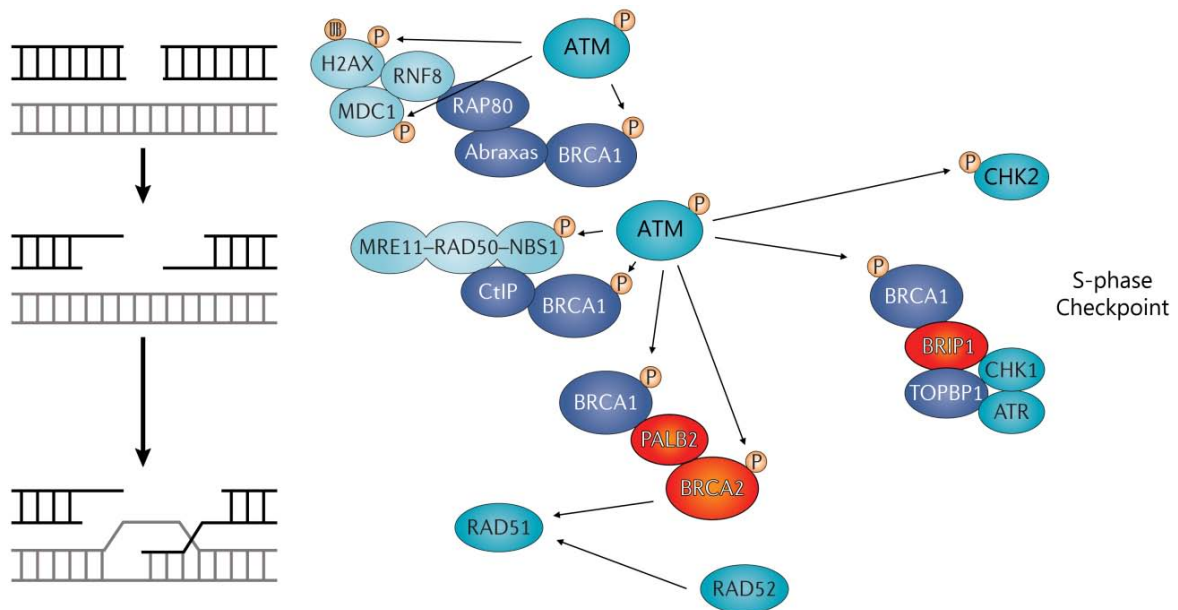


Figura 7: Ruta FA/BRCA y HRR: los macro complejos de BRCA1/FANCS y la proteínas FA (en naranja) involucradas en los pases de la HRR (adaptada de (Roy et al., 2012))

Las cortas colas de 3' ssDNA formadas entonces se alargan debido a una resección del DNA más amplia por parte de la exonucleasa 5'-3', EXO1 y de la acción concertada del complejo BLM-TOPOIII α -RMI1-RMI2 con la endonucleasa DNA2 (Mimitou and Symington, 2009, Symington, 2014). Los amplios trectos de ssDNA formados varían en longitud desde unos pocos cientos de nucleótidos a decenas de millares, dependiendo de la disponibilidad y la ubicación de la región homóloga en la cromátida hermana.

Por otro lado, el complejo MRN recluta también la proteína quinasa ATM facilitando los sucesivos pasos de la HRR. ATM reclutado al sitio de daño mantiene activa, potencia y transmite la señal de respuesta al daño a través de la fosforilación de efectores como las proteínas que regulan los puntos de control del ciclo celular, de los factores que remodelan la cromatina y de otros componentes de la reparación del DNA (Lavin, 2007, Lee and Paull, 2007, Lavin and Kozlov, 2007, Wu et al., 2012). El ssDNA es recubierto y estabilizado por RPA y el complejo ssDNA-RPA activa la proteína quinasa ATR y otro macro complejo, que comprende BRCA1/FANCS, BRIP1/FANCI y TOPBP1, promueve la actividad de ATR y de CHK1 para llevar a su completamiento la señalización para la puesta en marcha del punto de control de fase S (Greenberg et al., 2006, Burrows and Elledge, 2008). Entonces, otro

complejo compuesto de BRCA1/FANCS, PALB2/FANCD1 y BRCA2/FANCD1 promueve el intercambio entre RPA y RAD51/FANCR en el ssDNA (Benson et al., 1998, McIlwraith et al., 2000, McIlwraith and West, 2008). La proteína BRCA2/FANCD1 es imprescindible en este proceso y es uno de los mediadores claves de la HRR (Jensen et al., 2010, Zhang et al., 2009, Sy et al., 2009). La proteína DSS1 actúa mediando la interacción entre el complejo BRCA2/FANCD1-RAD51/FANCR y RPA para facilitar el intercambio entre RPA y RAD51/FANCR sobre del ssDNA. La DSS1 funciona disminuyendo el afinidad de RPA para el ssDNA y permitiendo así que RAD51/FANCR sustituya a RPA (Zhao et al., 2015). El complejo RAD51/FANCR-ssDNA entonces permite la invasión del filamento de ssDNA en la cromátida hermana para encontrar la región homóloga que proporciona la información genética correcta para la reparación (New et al., 1998, Shinohara and Ogawa, 1998, Sung, 1994, Bianco et al., 1998). Además de RAD51/FANCR, otro grupo de proteínas conocidas como “proteína parálogas de RAD51” juega un papel esencial en la HRR: las proteínas parálogas de RAD51 constan de cinco proteínas entre las cuales RAD51C/FANCO que son necesaria para que RAD51/FANCR pueda ser reclutado al sitio del daño y para la transducción de la señal de daño a las quinasas efectoras (Thacker, 2005, Suwaki et al., 2011). Se han identificado 5 proteínas parálogas de RAD51/FANCR en mamíferos y aves; la proteínas RAD51B, RAD51C/FANCO, RAD51D y XRCC2 forman un complejo denominado BCDX2 y otro complejo denominado CX3 está formado por RAD51C/FANCO y XRCC3 (Masson et al., 2001, Yonetani et al., 2005). Las proteínas parálogas de RAD51 estimulan la HRR modificando la estructura física del complejo RAD51/FANCR-ssDNA y favoreciendo la invasión de la cromátida hermana (Taylor et al., 2015) . Después de la etapa de invasión, la polimerasa eta (Pol η) alarga el filamento (McIlwraith et al., 2005) (Figura 6) y a partir de este punto hay dos vías posibles: una que produce la formación de uniones de Holliday (Figura 6A y B) y una que no (Figura 6C). Las uniones de Holliday son estructuras complejas que requieren para ser resueltas DNA nucleasas como GEN1, SLX1, la proteína FANCP/SLX4 (Ip et al., 2008, Fekairi et al., 2009, Svendsen et al., 2009) y DNA helicasas, como las helicasas Werner Syndrome (WRN) y BLM, ambas pertenecientes a la familia de las helicasas RecQ (Constantinou et al., 2000, Karow et al., 2000). Por contrario la vía alternativa llamada “Emparejamiento de hebra dependiente de síntesis” (*Synthesis Dependent*

Strand Annealing) puede completar la reparación del DSB sin entrecruzamiento cromosómico al no formarse una unión de Holliday (Nassif et al., 1994) (Figura 6C).

II.6. FA y la susceptibilidad al cáncer familiar de mama y ovario.

Los genes *BRCA2/FANCD1*, *BRCA1/FANCS*, *BRIP1/FANCI*, *PALB2/FANCN*, *RAD51C/FANCO* y *FANCM* son los genes más importantes de susceptibilidad al cáncer de mama y ovario (*Familiar Breast and Ovary Cancer*; FBOC) (Bogliolo and Surrallés, 2015). Cuando ambos alelos de los genes *BRCA2/FANCD1*, *BRIP1/FANCI*, *FANCN/PALP2*, *RAD51C/FANCO* y *BRCA1/FANCS* están mutados a nivel germinal, los pacientes sufren FA (*BRCA2/FANCD1*, *BRIP1/FANCI*, *FANCN/PALP2*) o una enfermedad FA-like (*BRCA1/FANCS* y *RAD51C/FANCO*) mientras que en los individuos con solo un alelo mutado, como es el caso de los padres de estos pacientes, se incrementa el riesgo de padecer un cáncer a lo largo de la vida (Levy-Lahad, 2010, D'Andrea, 2010) (Figura 8). Las mutaciones en *RAD51C/FANCO* influyen más la susceptibilidad al cáncer de ovario que de mama (Loveday et al., 2012, Osorio et al., 2012), y están asociadas con otros tipos de tumores como los HNSCC (Scheckenbach et al., 2014, Gresner et al., 2014). *RAD51C* y *FANCM* se asociaron con FA antes de ser también reconocidos como genes de susceptibilidad a FBOC (Meindl et al., 2010, Kiiski et al., 2014) y este hecho pone en evidencia la importancia del papel de la investigación sobre de la FA en oncología molecular. Históricamente solo los genes FA con una función por debajo de la monoubiquitinación del complejo ID y directamente involucrados en la HRR estaban relacionados con FBOCs (Levy-Lahad, 2010, D'Andrea, 2010). Sin embargo en los últimos años este modelo se ha puesto en discusión: los genes *FANCP/SLX4* y *FANCO/ERCC4* que actúan por debajo de la etapa de monoubiquitinación se han visto no tener relevancia para FBOC en las poblaciones Italiana, Alemana, Española, Estonia, y en la población Judía y no Judía Americana (Landwehr et al., 2011, Catucci et al., 2012, Fernandez-Rodriguez et al., 2012, Bakker et al., 2013, de Garibay et al., 2013, Shah et al., 2013, Osorio et al., 2013, Kohlhase et al., 2014), aunque *FANCP/SLX4*, por ejemplo, participe de forma substancial en la HRR por su papel

en la resolución de las uniones de Holliday (Fekairi et al., 2009, Munoz et al., 2009, Saito et al., 2009, Svendsen et al., 2009). Sin embargo variantes deletéreas en *FANCM* están asociadas con la susceptibilidad de los portadores al cáncer de mama en varias poblaciones (Kiiski et al., 2014, Peterlongo et al., 2015), lo que sugiere para *FANCM* un papel más complejo, independiente del FANCore y probablemente dependiente de su función en el reclutar el complejo Bloom (BLM-TOPOIII α -RMI1-RMI2) a las horquillas de replicación bloqueadas por los ICL (Deans and West, 2009). De hecho, a pesar de que *FANCM* no es esencial para la formación de los *foci* de RAD51 y para la HRR, las células FA-M comparten con las células FA-D1 y FA-N la sensibilidad a la Camptotecina, que inactivando la Topoisomerasa I, es un potente inductor de DSBs.

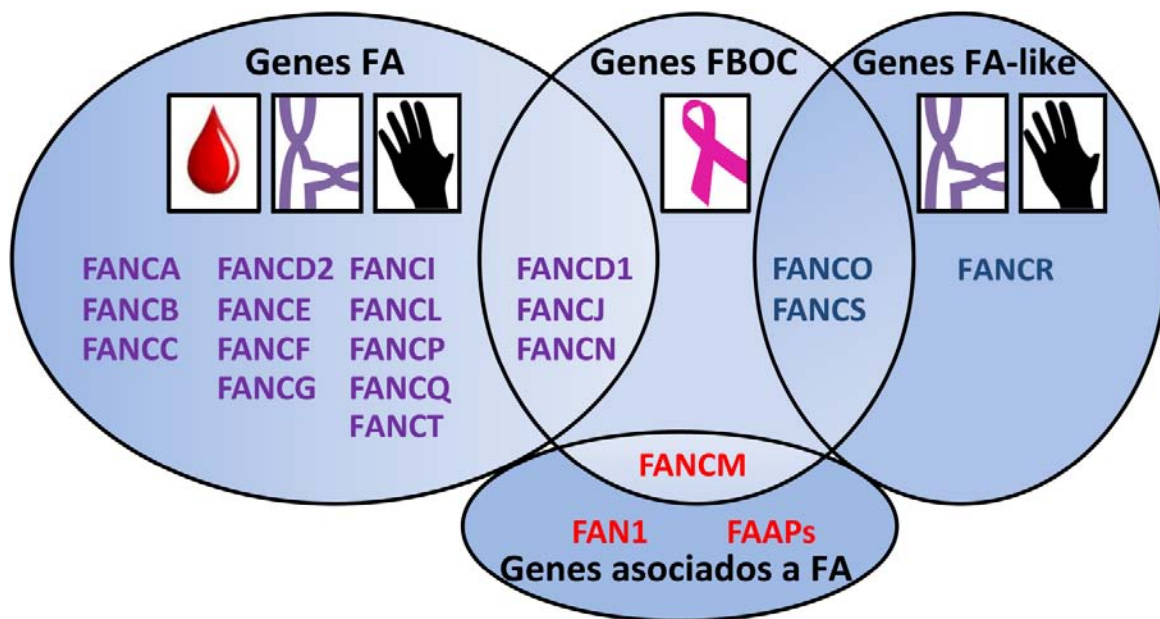


Figura 8: Mutaciones bialélicas en 15 genes (*FANCA, B, C, BRCA2/D1, D2, E, F, G, I, BRIP1/J, L, PALB2/N, SLX4/P, ERCC4/Q* y *UBE2T/T* en violeta) causan FA que se define por tres características fundamentales: BMF, fragilidad cromosómica y malformaciones. Mutaciones bialélicas en tres genes (*RAD51C/FANCO, BRIP1/J* y *RAD51/R* en azul) producen un síndrome FA-like sin BMF. Hay varios genes asociados a FA que pero no están relacionados con un fenotipo clínico FA (varios *FAAP, FANCM* y *FAN1*). Seis genes (*FANCD1, J, M, N, O,* and *S*) están asociados con cáncer familiar de mama y ovario subrayando la estricta relación entre FA y predisposición al cáncer en la población general (Modificada de (Bogliolo and Surrallés, 2015)).

II.6.1. Mutaciones bialélicas en genes FBOC: fenotipos celulares y clínicos

El fenotipo de las células FA-D1,-N, M, O y S es más complejo que el de las células que presentan mutaciones en genes para las proteínas del FANCore. Las células que presentan mutaciones bialélicas en genes FBOC, además que a los ICL, son sensibles a los agentes que producen un incremento de DSBs de forma directa como la radiación ionizante (IR) y la Camptotecina, o de forma indirecta como los inhibidores de la PARP (Tan et al., 2008). En estas células (pero no en las FA-M) también se puede apreciar una disminución consistente de la formación de los *foci* de RAD51 y de la HRR. El fenotipo clínico de los pacientes FA con mutaciones en *BRCA2/FANCD1* es más severo que el fenotipo FA clásico, con una propensión a desarrollar BMF, tumores sólidos y leucemias en edades mucho más tempranas que las de los pacientes que tienen mutaciones en los genes FA involucrados en las etapas anteriores a la activación del complejo ID (Wagner et al., 2004). Los pacientes FA asignados al subtipo FA-N, cuyas células contienen mutaciones bialélicas en *PALB2*, también empiezan muy pronto a lo largo de la vida a desarrollar cánceres. La proteína PALB2 se une directamente a la proteína BRCA2/FANCD1 permitiendo así su función en la HRR (Oliver et al., 2009), y esto puede explicar la clínica muy parecida de los pacientes FA-N y FA-D1. Mutaciones en el gen *FANCO/RAD51* y *BRCA1/FANCS* causan un síndrome FA-like (Vaz et al., 2010, Sawyer et al., 2014) y las células que derivan de individuos con mutaciones en estos genes son sensibles a la Camptotecina y tienen niveles reducidos de *foci* de RAD51 después de daño al DNA de forma muy parecida a las células deficientes para *BRCA2* y *PALB2*. Sin embargo la ausencia de BMF es en contraste con el fenotipo de los pacientes FA-D1 y FA-N (Reid et al., 2007, Alter et al., 2007) posiblemente debido al hecho que las mutaciones presentes en los pacientes podrían tener un carácter hipomórfico. Sin embargo, la deficiencia de BRIP1/FANCI, aunque interactúe con BRCA1/FANCS, proteína fundamental para la HRR y la resistencia a los DSBs (Venkitaraman, 2002), produce en las células un fenotipo más parecido a lo que confiere la falta de una cualquiera de las proteínas que ejercen sus funciones en las etapas previas a la activación del complejo ID. Las células con mutaciones bialélicas para *BRIP1/FANCI* no tienen sensibilidad ni a la Camptotecina ni a la IR y forman *foci* normales de RAD51 después de daño al DNA. Se postula por lo tanto que BRIP1/FANCI pueda tener un papel distinto en la reparación del DNA respecto a la

etapa de la HRR (Bridge et al., 2005, Singh et al., 2009, Cantor and Andreassen, 2006) posiblemente relacionado con su actividad en resolver estructuras complejas del DNA como las estructuras 4G capaces de bloquear las horquillas de replicación (Castillo Bosch et al., 2014). Pero, sin embargo, *BRIP1/FANCF* es un gen de susceptibilidad al cáncer de mama mientras que *FANCP/SLX4* no lo es no obstante tenga un papel importante en la resistencia a la Camptotecina, a la IR (Munoz et al., 2009) y a los inhibidores de la PARP (Shah et al., 2013, Bakker et al., 2012, Catucci et al., 2012, Bakker et al., 2013, de Garibay et al., 2013). Por lo tanto, en falta de un ensayo que nos permita discriminar, todos los nuevos genes FA, FA-like o FAAP que se vayan identificando se deberían también considerar por defecto candidatos para la susceptibilidad al cáncer de mama y/o de ovario.

II.7. La secuenciación del DNA de nueva generación: secuenciación del exoma

En las últimas tres décadas, ha habido enormes progresos en la identificación de los grupos de complementación y de los genes involucrados en la FA. Diferentes estrategias como la clonación por complementación, la co-IP acoplada a espectrometría de masa y estudios de tipo más clásico como la MMCT o la análisis del ligamiento con marcadores polimórficos seguida de clonaje posicional se han utilizado con éxito para encontrar los genes alterados en los pacientes FA (de Winter and Joenje, 2009). También el mapeo por array de SNPs de regiones del genoma en homocigosis se ha demostrado exitoso en el caso del descubrimiento de *RAD51C/FANCO* gracias a la consanguinidad de la familia analizada (Vaz et al., 2010). Sin embargo, hay algunos pacientes dentro de la población FA que aún no se habían podido caracterizar. Los grupos de complementación de estos pacientes no se pudieron determinar porque todas las técnicas utilizadas para la identificación de los genes mutados no produjeron ningún resultado positivo. Las técnicas tradicionales son ineficaces en estos pocos pacientes por ser los grupos de complementación extremadamente raros y con un limitado número de casos no emparentados provenientes de familias no informativas. Sin embargo, las técnicas de secuenciación de nueva generación (*Next Generation Sequencing*; NGS) como la secuenciación del exoma (*Whole Exome Sequencing*; WES) o la

secuenciación de genoma (*Whole Genome Sequencing*; WGS) ofrecieron nuevas oportunidades para dilucidar la base genética de las enfermedades extremadamente raras. Esta prueba de concepto fue demostrada con el síndrome de Freeman-Sheldon. El DNA de cuatro pacientes no emparentados fue sometido a WES y *MYH3* se identificó como el gen candidato único. Las variantes encontradas en *MYH3* no habían sido previamente identificadas, es decir, no estaban catalogadas como SNPs ni presentes en los datos de WES de la base de datos HapMap. Este estudio demostró la viabilidad de la aplicación de la WES para identificar el gen candidato para un trastorno mendeliano, a pesar de solo tener un pequeño número de casos no emparentados (Ng et al. 2009).

En el genoma humano hay alrededor de 180.000 exones: estos constituyen aproximadamente el 1% del genoma, que se traduce en alrededor de 30 megabases de longitud (Ng et al., 2009). Por lo tanto, en comparación con la WGS, el análisis del exoma requiere menores cantidades de datos para lograr la cobertura deseada para la detección de variantes. Se estima que las alteraciones en las regiones del genoma humano codificantes para las proteínas constituyen la causa de aproximadamente el 85% de las enfermedades producidas por mutaciones (Choi et al., 2009). La principal limitación de los estudios de WES es que, hasta ahora, se han centrado principalmente en los exones y en los sitios de splicing con poca o ninguna cobertura de otras secuencias como los intrones, los promotores, los potenciadores, los microRNAs y los otros elementos reguladores (Ku et al., 2011, Metzker, 2010).

II.7.1. WES: Captura y amplificación de las regiones codificantes.

La WES implica, antes de la secuenciación, la captura de las regiones codificantes del DNA genómico que se quiere analizar. Todos los métodos de WES comienzan con romper (enzimáticamente o por sonicación) al azar el DNA genómico para conseguir fragmentos de un determinado tamaño que se puedan purificar, por ejemplo, en un gel de poliacrilamida recortando una banda en correspondencia del tamaño de interés. Se crea así una biblioteca de fragmentos cuyos extremos se ligan a secuencias adaptadoras complementarias a cebadores universales. A las secuencias adaptadoras se pueden añadir sucesivamente por PCR

secuencias específicas de DNA que funcionan como un código de barras identificando de forma inequívoca a las librerías. Los códigos de barras reducen drásticamente los costes al permitir al usuario agrupar librerías de diferentes muestras en una única reacción de secuenciación. El software de análisis se encargará sucesivamente de reconocer y ensamblar las diferentes secuencias en función de los diferentes códigos de barra. El sistema Nextera[®] de Illumina[®] utiliza una variante hiperactiva de la Transposasa Tn5 para catalizar a la vez la fragmentación del DNA y la incorporación de los adaptadores en una única reacción reduciendo drásticamente los tiempo de preparación de las librería (Adey et al., 2010). Una vez que la librería esté lista, se procede a la captura de las secuencias exónicas. Una de las técnicas de enriquecimiento más utilizada es la captura en fase sólida. Los fragmentos de DNA se desnaturalizan y se hibridan sobre de microarrays con secuencias de nucleótidos específicas para las secuencias exónicas (Albert et al., 2007, Hodges et al., 2007). Los fragmentos que no se han hibridado son eliminados por lavado. Otra técnica de enriquecimiento, esta vez en fase líquida, se basa en la hibridación de los fragmentos con una librería de sondas para el exoma marcadas con biotina. Sucesivamente, utilizando bolas magnéticas recubiertas de estreptavidina, se purifican las secuencias que corresponden a los exones (Gnirke et al., 2009).

Después de la captura, el DNA se vuelve a desnaturalizar y los fragmentos de la librería se suelen amplificar por PCR en emulsión (emPCR) o por PCR en fase sólida (sfPCR) (Figura 9). La emPCR permite la amplificación de una sola molécula de DNA a partir de una emulsión de agua en aceite. En la emulsión el agua forma gotitas estables, incluso a altas temperaturas, que funcionan como micro-reactores (Nakano et al., 2003). Las gotitas en la emulsión tienen un diámetro de alrededor de 10 μm y contienen los cebadores, la mezcla para la reacción de PCR y (idealmente) una sola molécula de DNA molde (Figura 9A). De esta forma los fragmentos de la librería se amplifican de forma individual, homogénea y sin interferencias. En las emPCR para NGS uno de los cebadores está anclado a una bolita de 1 μm y de esta forma los productos finales de la amplificación quedarán ligados a un soporte sólido (Figura 9A). Tras el éxito de la amplificación, millones de estas micro-esferas se pueden distribuir e inmovilizar sobre de un soporte físico que puede ser distinto

dependiendo de la tecnología que se utilizará para secuenciar (Figura 9A) (Shendure et al., 2005, Kim et al., 2007b, Leamon et al., 2003).

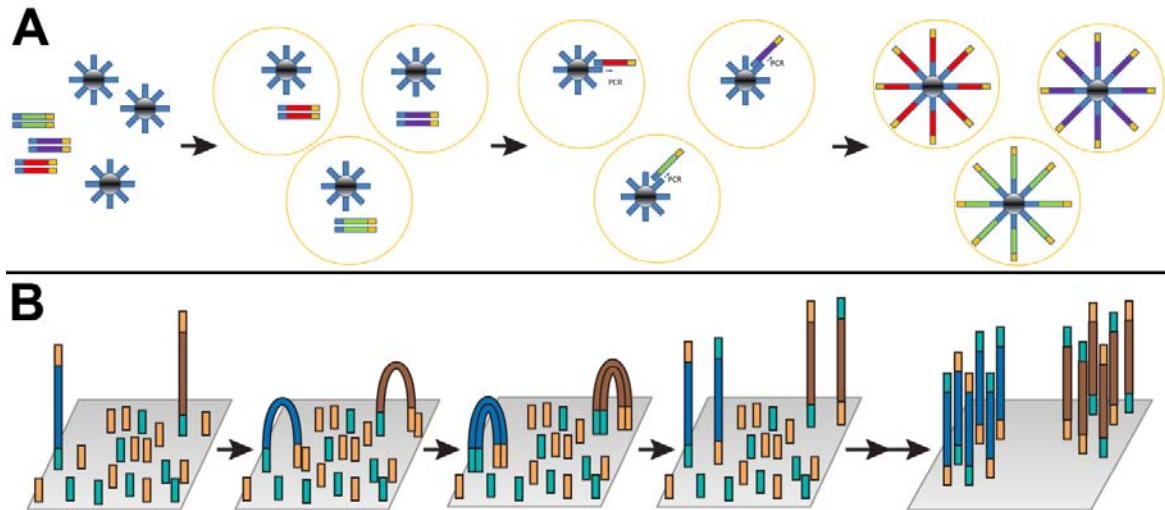


Figura 9: Amplificación de las librerías mediante PCR en emulsión (A) y PCR en fase sólida (B). Explicación en el texto (adaptada de (Shendure and Ji, 2008)).

La sfPCR es al día de hoy la técnica más utilizada para amplificar librerías de DNA para NGS. La sfPCR se utiliza para producir “colonias” de moléculas amplificadas de forma clónica distribuidas al azar sobre de un soporte (Figura 9B). Ambos cebadores utilizados para la sfPCR se encuentran en alta densidad atados covalentemente por un extremo a la superficie del soporte (Figura 9B) y la proporción entre los cebadores y el DNA patrón determina la densidad de las colonias de moléculas amplificadas. El fragmento de DNA de cadena sencilla de la librería cae en este “césped” de cebadores y se ata a uno de ellos a través de la secuencia adaptadora. Mediante una reacción de extensión se crea una molécula de DNA de doble cadena con una de las dos hebras atadas ahora covalentemente al soporte. Se desnaturaliza el DNA eliminando la cadena libre y se consigue que el DNA a cadena sencilla se encuentre ahora ligado covalentemente al soporte y listo para empezar la amplificación. La amplificación de los fragmentos procede por formación de “puentes” de DNA de doble cadena donde cada una de las cadenas está anclada al soporte a través del primer (Figura 9B). Una sfPCR puede producir 100–200 millones de “colonias” de

moléculas amplificadas con una extremidad libre (Adessi et al., 2000). Terminada la sfPCR, las cadenas complementarias que se han producido en el proceso de amplificación se cortan y se eliminan antes del comienzo de la secuenciación. Se añade el cebador para secuenciar junto con la mezcla de los nucleótidos fluorescentes y la polimerasa y la reacción de secuenciación puede empezar (Figura 9B).

Paralelamente a estos sistemas, también se han desarrollado sistemas que permiten secuenciar moléculas únicas de DNA sin la necesidad de una etapa previa de amplificación. En estos sistemas la preparación de la muestra es enormemente simplificada al no necesitarse ni grandes cantidades ni la amplificación del DNA. Cada manipulación de la muestra, y especialmente la amplificación, pueden causar artefactos cuantitativos y cualitativos (Dohm et al., 2008, Milos, 2009). El método más utilizado actualmente para secuenciar moléculas únicas de DNA es la secuenciación en tiempo real (*Single Molecule Real Time Sequencing*; SMRT) puesto a punto por Pacific Biosciences[®] (Korlach et al., 2010). La preparación de las muestras para el análisis con esta tecnología supone una incubación de los fragmentos de las librerías para secuenciar, ya emparejados con los cebadores, con moléculas de DNA polimerasa marcadas con una cola de biotina. Se forman así complejos estables DNA/polimerasa que, antes de empezar la reacción de secuenciación, se enganchan al fondo de unos nano-pocillos recubiertos de estreptavidina (Eid et al., 2009)

II.7.2. Secuenciación de las librerías

Una vez que el DNA esté listo y atado a un soporte se procede a la secuenciación y a la adquisición de la información. Como en el caso de la preparación del DNA también para esta etapa hay diferentes aproximaciones. Una de estas es la terminación cíclica reversible (CRT) que consiste en utilizar nucleótidos marcados con fluorocromos en el 3'OH que, una vez incorporados en el DNA, bloquean la DNA polimerasa. En una primera etapa se añade la mezcla que contiene el cebador universal, los cuatro nucleótidos fluorescentes y la polimerasa que incorpora un nucleótido y después se bloquea al no tener un extremo 3'OH libre para seguir polimerizando el DNA. Se eliminan los nucleótidos no incorporados y se excita el fluorocromo y se toma una imagen digital. Entonces se elimina químicamente el

fluorocromo liberando el extremo 3'OH y se vuelven a añadir los cuatro nucleótidos fluorescentes permitiendo así a la polimerasa de poder incorporar uno nuevo (Mitra et al., 2003) (Figura 10).

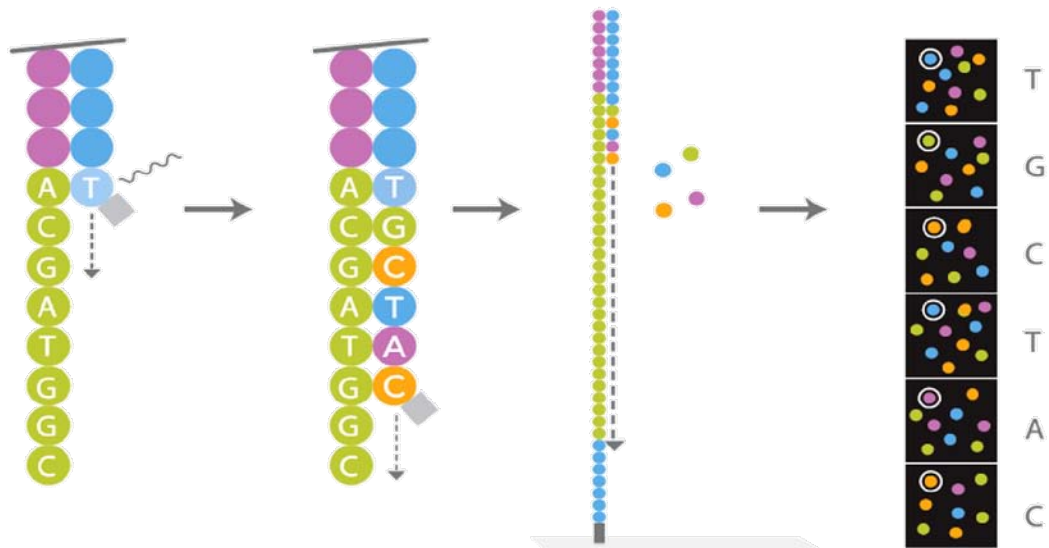


Figura 10: Secuenciación por terminación cíclica reversible: explicación en el texto (modificada de: <http://www.cegat.de/>)

El proceso se repite igual hasta terminar la reacción de secuenciación. En el caso del DNA amplificado en fase solida (Illumina®) es posible secuenciar también desde el otro extremo (*Paired-end sequencing*) consiguiendo así la secuencia complementaria del fragmento. Para conseguirlo se desnaturaliza el DNA eliminando la cadena complementaria que se ha formado a lo largo de la primera secuenciación, se vuelve a pasar por una fase de sfPCR pero, antes de proceder a la nueva secuenciación, esta vez es la cadena molde la que viene eliminada. Este proceso permite generar datos de secuenciación de mejor calidad facilitando la alineación con la secuencia de referencia y aumentando la capacidad del sistema para detectar correctamente deleciones, inserciones y secuencias repetidas. Los sistemas de Illumina llegan a secuenciar fragmentos de hasta 300 nucleótidos y actualmente esta tecnología es la más utilizada para las aplicaciones de NGS y la con más soporte técnico y de aplicaciones del mercado (van Dijk et al., 2014).

Otra aproximación es la secuenciación por ligación (*Sequencing by ligation*; SBL) que utiliza una ligasa en lugar de una DNA polimerasa para la secuenciación del DNA y es por tanto un sistema conceptualmente muy distinto respecto a la CRT o a la SMRT. La secuenciación se organiza en fases distintas llamadas “tandas de cebadores” y cada tanda consiste en la hibridación de un cebador seguida por una serie de ciclos de hibridación de sondas (Figura 11A).

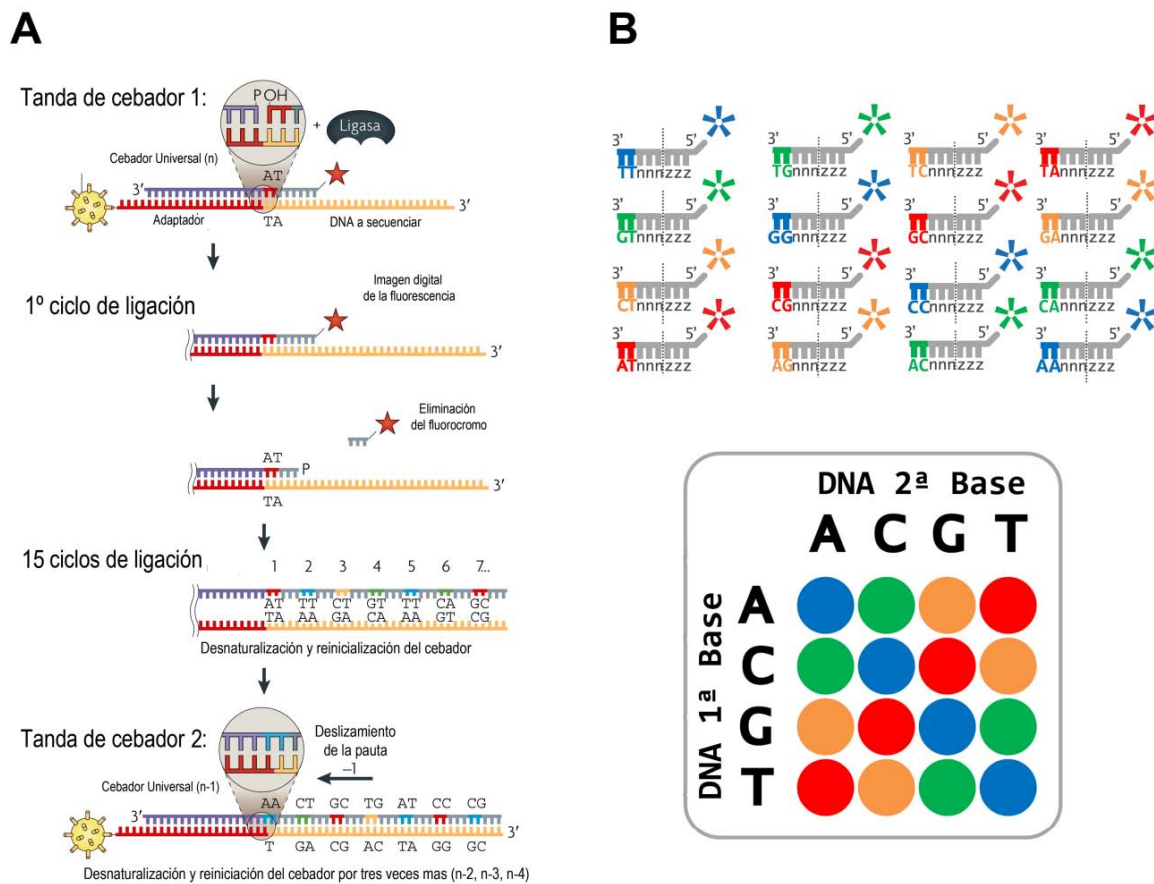


Figura 11: (A) Tanda de secuenciación con el sistema Solid™: después de la hibridación con el cebador, las 16 sondas diferentes se añaden en el sistema. Las que encuentran las bases complementarias son atadas por una ligasa. Una vez adquirida la imagen digital se cortan las 3 últimas bases eliminando el fluorocromo y dejando un extremo 5' fosforilado donde se puede atar la sonda del ciclo sucesivo. (B) Combinaciones de colores para las 16 posibles sondas que se utilizan en el sistema Solid™ (adaptada de (Metzker, 2010)).

Las sondas están constituidas por dos bases normales, tres bases degeneradas y tres bases universales y la última base lleva atado un fluorocromo. Hay 16 posibles combinaciones de dos bases pero solo 4 colores y por lo tanto habrá más de una combinación

de bases con el mismo color según el esquema en (Figura 11B). Por cada ciclo una sonda se empalma a las bases complementarias en el DNA gracias a sus dos bases normales y una ligasa la une a la sonda precedente (o al cebador si es el primer ciclo). Después de haber eliminado por lavado las sondas que no se hibridaron, se excita el fluorocromo y se adquiere una imagen digital en colores. Entonces se eliminan las tres bases universales y el fluorocromo y se procede a un nuevo ciclo de hibridación (Shendure et al., 2005). Una vez completados 15 ciclos de hibridación de sondas, el sistema se reinicializa por desnaturalización dejando así el fragmento de DNA disponible para un nuevo cebador que se hibridará desplazado una base con respecto al cebador anterior. Por lo tanto, en la primera tanda del proceso se determinan las posibles identidades de las bases en las posiciones 4, 5, 9, 10, 14, 15, etc. mientras que en la segunda tanda, con el cebador desplazado una base, el proceso determina las posibles identidades de las bases en posición 3, 4, 8, 9, 13, 14, etc. Las tandas se repiten hasta llegar a la posición 0, la última base del adaptador. La característica principal de este sistema es que cada base se “lee” dos veces en reacciones independientes, aumentando así la fidelidad de la secuenciación. La secuencia de colores o *color space* obtenida después de las 5 tandas de cebadores se ordena linealmente y se compara con el *color space* del genoma de referencia para descodificar la secuencia del DNA (Figura 12). Con el sistema Solid™ también es posible secuenciar desde el otro extremo del fragmento utilizando sondas que tienen las dos bases normales a partir del 5' y los fluorocromos en 3' aunque a causa de las reacciones químicas diferentes es posible secuenciar de forma fiable solo por tramos de hasta 35 nucleótidos. Las principales limitaciones de este sistema son la imposibilidad de conseguir secuencias más larga de 75 nucleótidos y de no poder hacer secuenciación *de novo* cuando no se conoce la secuencia de referencia. Sin embargo, esta última limitación se resolvió en los últimos años con el desarrollo de un sexto cebador y una mezcla única de sondas para utilizarse después de las primeras cinco tandas. Esta tecnología, permite de identificar las bases discretas sin un genoma de referencia gracias un complejo algoritmo matemático (Massingham and Goldman, 2012) haciendo posible así la secuenciación *de novo* con el sistema Solid™.

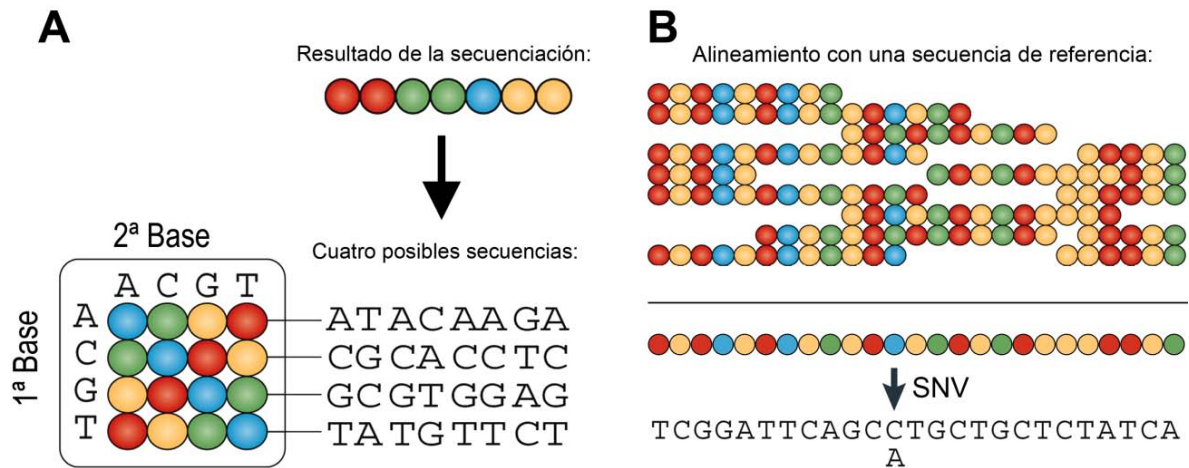


Figura 12: (A) Descodificación de datos de SOLiD™: el resultado de la secuenciación es una combinación de colores correspondiente a cuatro posibles secuencias de DNA. En principio las combinaciones de colores se pueden transformar en la secuencia correcta si conocemos por lo menos una de las bases de la secuencia. (B) En realidad los datos son inmediatamente alineados de forma automática a una secuencia de referencia por el software de análisis. Las variantes se identifican por los cambios de dos colores contiguos con respecto a la referencia. Cuando cambia solo un color es por un error del proceso de secuenciación (adaptada de (Metzker, 2010)).

II.7.3. Otras tecnologías para secuenciación de librerías

El Ion Torrent es un método de secuenciación de DNA basado en la detección de iones de hidrógeno que se liberan durante la polimerización del DNA. Unos micro pocillos, que contienen cada uno una bolita con atado el DNA que se ha amplificado por emPCR, se llenan con una solución que lleva una sola especie de desoxirribonucleótido trifosfato (dNTP) junto a la polimerasa. Si el dNTP introducido es complementario al DNA, se incorpora y esto provoca la liberación de un ion de hidrógeno que activa un sensor puesto de bajo del pocillo. Si hay una serie de bases que se repiten iguales, múltiples dNTPs se incorporaran en un solo ciclo liberando cada uno un ion de hidrogeno y el sensor registrará una señal proporcionalmente mayor (Rothberg et al., 2011). Se eliminan entonces los nucleótidos no incorporados y se introduce en el sistema otro nucleótido diferente. Esta tecnología de secuenciación se diferencia por no utilizar nucleótidos modificados ni ópticas para tomar imágenes digitales y esto hace que sea la alternativa más económica para la NGS. Los sistemas Ion Torrent pueden llegar a secuenciar fragmentos de hasta 400 nucleótidos pero son propensos a cometer errores en tramos de DNA donde se repiten muchas veces los

mismos nucleótidos (van Dijk et al., 2014) y no pueden secuenciar la cadena complementaria. Parecida a la tecnología Ion Torrent es la tecnología de piro-secuenciación utilizada por Roche que en lugar de detectar los iones de hidrogeno detecta el pirofosfato que se libera al formarse el enlace entre los nucleótidos en la síntesis del DNA. (van Dijk et al., 2014).

La tecnología SMRT aprovecha las ventajas derivadas de la evolución natural de la DNA polimerasa utilizándola para secuenciar el DNA en tiempo real y sin interrupción. Para superar los retos inherentes a la observación de una enzima de 15 nm, se utilizan placas con miles de nano-pocillos que tienen un diámetro de pocas decenas de nanómetros (*zero-mode waveguides*; ZMW). El diámetro de estos pocillos es mucho menor que la longitud de onda de la luz de excitación y, por lo tanto, la luz penetra solamente unos pocos nanómetros dentro del pocillo. Se crea así un volumen de detección que permite observar individualmente a las moléculas fluorescentes. Dentro de cada ZMW está un único complejo DNA/polimerasa atado al fondo del pocillo de tal manera que reside de forma permanente dentro del volumen de detección. En este volumen, la incorporación de nucleótidos marcados con fluorescencia en una cadena de DNA se puede detectar en medio del ruido de fondo porque el volumen de detección está ocupado por los nucleótidos libres solo una pequeña fracción de tiempo siendo el tiempo de difusión de unos pocos microsegundos. A medida que se incorporan los nucleótidos, cada base queda retenida en el sitio catalítico de la polimerasa (y por lo tanto dentro del volumen de detección) por decenas de milisegundos, un tiempo de decenas de miles de veces más largo que el tiempo de difusión. Durante este período, el fluorocromo unido a la polimerasa se excita y emite luz fluorescente del color correspondiente a la identidad de la base. Entonces la polimerasa, como en el proceso natural de la síntesis del DNA, escinde el enlace fosfato que conecta el nucleótido al fluorocromo, permitiendo que el fluorocromo se difunda rápidamente fuera del volumen de detección. (Esta aproximación tiene la capacidad de producir secuencias individuales y continuas hasta 20.000 nucleótidos de longitud (Eid et al., 2009, Utturkar et al., 2014) pero es en absoluto la más costosa del mercado.

II.7.4. NGS y diagnóstico de la FA

Actualmente, después de un diagnóstico de FA con el test de fragilidad cromosómica, para identificar el gen mutado en una paciente se sigue un proceso de trabajo complicado que puede requerir muchos meses para completarse y que necesita personal altamente especializado (Figura 13).

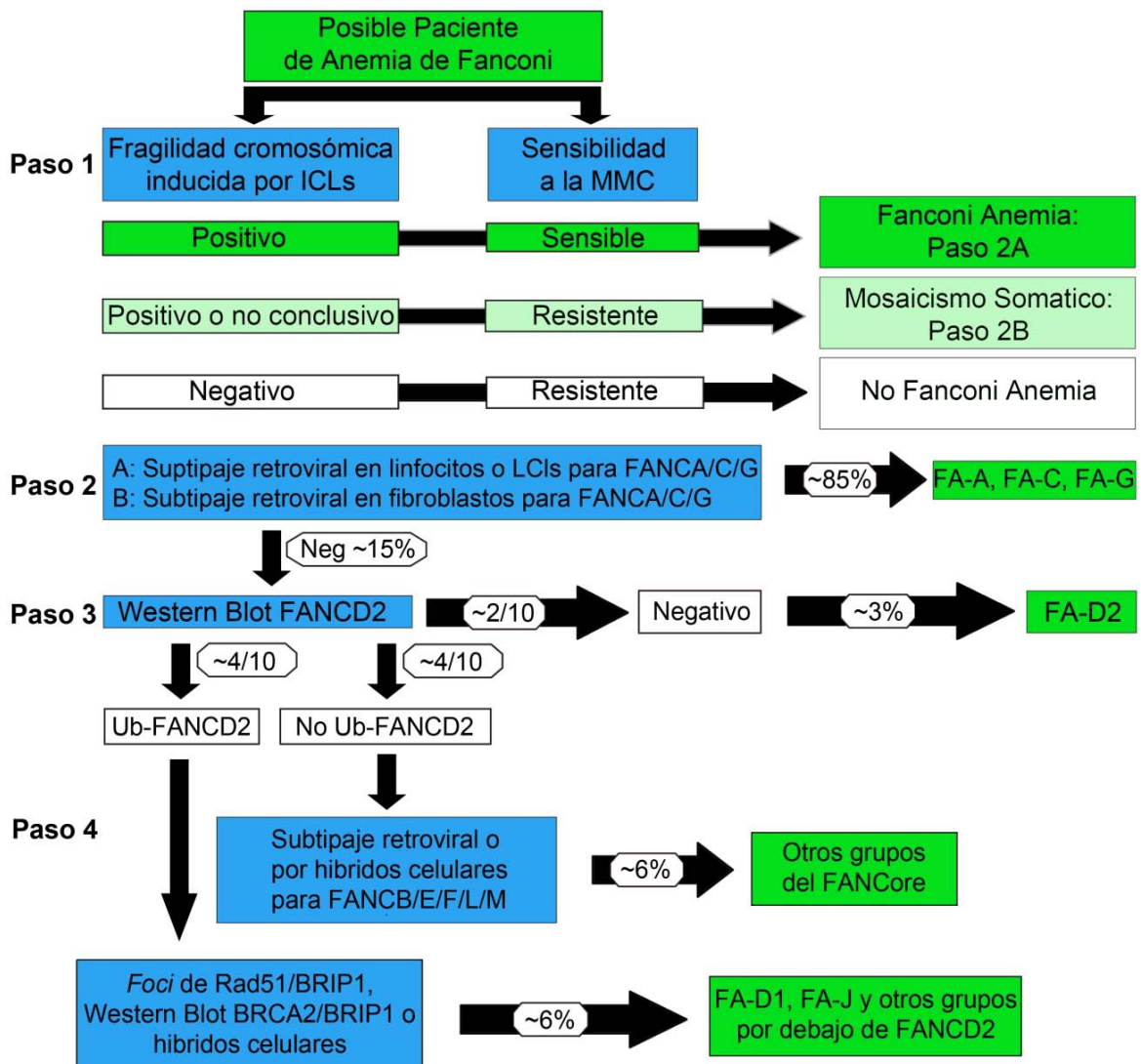


Figura 13: Esquema de los diferentes procesos para la confirmación del diagnóstico y la caracterización molecular de los pacientes de FA (Adaptada de (Casado et al., 2007))

El subtipaje genético por complementación se realiza mediante el análisis de la reversión de los fenotipos celulares de sensibilidad a la MMC y bloqueo del ciclo celular en fase G2/M después de introducir mediante vectores virales una copia salvaje de los genes FA en los linfocitos del paciente. Se suelen testar por primeros los genes *FANCA/C/G* debido a que se cree que son los más frecuentemente mutados en la población mundial (Auerbach, 2009). Si no hay éxito, entonces hay que establecer una línea celular linfoblastoide (LCL) y analizar por WB la expresión y la monoubiquitinación de FANCD2. Si la banda correspondiente a la proteína FANCD2 está ausente se puede adscribir el paciente al subtipo FA-D2. Si la banda de FANCD2 está presente pero no se monoubiquitina en respuesta a daño al DNA, el defecto estará en los genes restantes del FANCore y por lo tanto se vuelve a buscar el subtipo por transducción retroviral o fusión celular para los genes del complejo que quedan por testar. Si la banda de FANCD2 está presente y se monoubiquitina, el defecto estará en los genes que se encuentran por debajo de la monoubiquitinación de FANCD2 (Hanenberg et al., 2002, Casado et al., 2007) (Figura 13). Una vez identificado el subtipo del paciente, hay que caracterizar las mutaciones y esta etapa presenta también muchas dificultades. Algunos de los genes son muy grandes como por ejemplo el gen *FANCA* que tiene 43 exones y una *Open Reading Frame* (ORF) de 4368 nucleótidos o *BRCA2* con 26 exones y una ORF de 10257 nucleótidos. Además las mutaciones suelen estar distribuidas a lo largo de los genes sin claros “puntos calientes” y es muy común la presencia de mutaciones “privadas” o sea características solo de una familia y no reportadas en las bases de datos (Levrán et al., 2005b, Yagasaki et al., 2004, Savino et al., 2003). Esto no es cierto para algunos grupos minoritarios altamente consanguíneos dónde algunas mutaciones determinadas resultaron fijadas en la población por efecto fundador como en el caso de la etnia gitana española, (Callen et al., 2005) o la etnia Judía Ashkenazi, (Whitney et al., 1993) o la población Afrikaner de Sudáfrica (Tipping et al., 2001). A causa de esta extrema heterogeneidad mutacional se desarrollaron técnicas alternativas a la secuenciación clásica del DNA para encontrar las mutaciones por lo menos en los genes FA más frecuentemente mutados. La detección de mutaciones puntuales del gen *FANCA* (mutado entre un 60 y un 80% de los pacientes FA (Castella et al., 2011b)) por ejemplo, se realiza mediante cromatografía líquida desnaturizante de alto rendimiento (*Denaturing High-Performance Liquid Chromatography*; DHPLC). Se amplifican todos los exones del gen, incorporando

parte de las secuencias intrónicas a los dos lados y los productos obtenidos se analizan mediante DHPLC y se procede a la secuenciación directa del DNA solo de los fragmentos que aparecen tener mutaciones (Rischewski and Schneppenheim, 2001). Esta técnica pero tiene un problema para detectar las grandes deleciones que constituyen un 15-20% de todas las mutaciones en *FANCA* y por lo tanto se suele acoplar el análisis por DHPLC con el análisis por MLPA (*Multiplex ligation-dependent probe amplification*) (Schouten et al., 2002, Gille et al., 2012). La técnica de la DHPLC acoplada con la MLPA se ha desarrollado solo para detectar mutaciones en *FANCA*, *BRCA2/FANCD1* y *PALB2/FANCN* (<https://www.mrc-holland.com>), los genes *FANCD2* y *-I*, son analizados a partir del cDNA, y por lo tanto para estos pacientes se requiere el establecimiento de una línea celular, mientras que para todos los otros genes no hay un método de análisis consensuado.

A lo largo los últimos años, el rápido desarrollo de las tecnologías de NGS ha contribuido a reducir radicalmente el coste y el tiempo requerido para la WES y la WGS. El Proyecto Genoma Humano necesitó aproximadamente 13 años y costó varios miles de millones de dólares; al día de hoy, el mismo proceso se puede completar en cuestión de semanas para unos pocos miles de dólares usando NGS. Recientemente Illumina® ha realizado un nuevo sistema (HiSeq X Ten) que pretende abatir la barrera del “Genoma a 1000 dólares”. Si esto se confirmara significaría que el coste por la secuenciación de un genoma humano se abarató más de 10.000 veces respecto al 2004 (van Dijk et al., 2014). Esta caída en picado de los precios hace que ya se pueda plantear el uso cotidiano de las tecnologías de NGS no solo como herramienta para la investigación de base sino también en el diagnóstico de las enfermedades genéticas raras y heterogéneas como la FA. Cuando empezamos nuestra tesis doctoral en el 2010 el coste de la WES rodeaba los 7000 € por paciente mientras que ahora hay compañías que ofrecen WES por alrededor de 500€ -1000 € según se incluye o no el análisis bioinformático. En el caso de la FA el uso de la WES para la caracterización de los pacientes se presenta al día de hoy como una alternativa muy válida tanto para el subtipaje como para la caracterización de las mutaciones de todos los enfermos de FA independientemente del grupo de complementación. Utilizando la WES se pueden generar rápidamente informaciones muy importantes para los pacientes: la caracterización de las mutaciones, desde un punto de vista clínico, permite el diagnóstico prenatal y preimplantacional y realizar análisis de portadores en los familiares de los pacientes FA.

También permite descartar la enfermedad en donantes de medula ósea emparentados. Igualmente, las mutaciones pueden ser usadas como marcadores para comprobar el éxito de un trasplante de medula en los pacientes y para el desarrollo de nuevas terapias como la terapia génica (Tolar et al., 2012) o las más novedosas técnicas de *gene editing* (Rio et al., 2014, Genovese et al., 2014).

Hay pero que tener en cuenta las limitaciones de la WES: la WES tradicional no puede detectar grandes deleciones y por lo tanto en un 15-20% de los pacientes FA-A no se podrán encontrar todas las mutaciones. Este problema se podría parcialmente resolver si la cobertura del gen en cuestión es alta: en las regiones de hemicigosis la cobertura (o sea el número medio de veces que una región se ha “leído” en secuencias independientes) debería reducirse del 50% con respecto a un control donde son presentes los dos alelos mientras que en las regiones con deleciones en homocigosis la cobertura debería bajar a 0. Se están desarrollando programas bioinformático y algoritmos para detectar grandes deleciones a partir de los datos de NGS pero actualmente no hay todavía herramientas destinadas a usuarios sin grandes conocimientos bioinformáticos y, además, se ha visto que la mayoría de estas herramientas tiene todavía limitaciones importantes a la hora de analizar datos reales (Tan et al., 2014). A parte de las grandes deleciones, hay que tener también en cuenta el mosaicismo somático presente en entre un 15-20% de todos los pacientes FA, sobre todo a la hora de elegir la fuente del DNA para la WES. A nivel molecular, los mecanismos que pueden generar la reversión de la mutación como la conversión génica, la recombinación intragénica o la reversión directa de la mutación pueden llevar a la completa desaparición de uno de los alelos mutados en la medula de los pacientes (Lo Ten Foe et al., 1997, Kalb et al., 2006). El mosaicismo se suele poder detectar en células de sangre periférica (*peripheral blood*; PB) en el ensayo de fragilidad cromosómica porque después del tratamiento se observan dos poblaciones de células, una con varias roturas cromatídicas en cada célula y otra población sin roturas que corresponde a la población de células revertidas pero en algunos pacientes el nivel de mosaicismo puede ser tan alto que el test de fragilidad en linfocitos puede resultar del todo negativo. En estos pacientes se suele proceder al test en los fibroblastos primarios de piel porque hasta ahora no se han encontrado evidencias que el mosaicismo somático pueda ocurrir en este tipo de células (Pinto et al., 2009, Waisfisz et al., 1999, Lo Ten Foe et al., 1997). El utilización de los fibroblastos primarios de piel conlleva

pero el problema de tener que proceder a una biopsia en los pacientes, un proceso más invasivo que no siempre se acepta con facilidad. Por lo tanto, dependiendo de los niveles de mosaicismo, una de las dos mutaciones puede no ser detectable por WES en el DNA proveniente de PB de los pacientes mosaico. Una alternativa a la WES en el diagnóstico de las enfermedades genéticas complejas como la FA, es la NGS de paneles de genes: en el caso de la FA se pueden desarrollar kit de captura únicamente para la secuenciación de los 16 genes involucrados en la enfermedad evitando así de tener que secuenciar todo el exoma. Esta alternativa puede llegar a ser bastante más barata que la WES tradicional, sobre todo si acoplada con un secuenciador que se base en el sistema Ion Torrent (Nicchia et al., 2015). Los principales obstáculos a la utilización de los paneles de genes son el coste en términos económicos y de tiempo de la puesta a punto de las librerías y la necesidad de acumular un número suficiente de muestras (nuestro laboratorio confirma el diagnóstico de media de una decena de nuevos caso de FA al año) para amortiguar el coste de los reactivos. Ulteriores complicaciones pueden derivar de la falta de uniformidad en los datos generados cuando diferentes laboratorios utilizan diferentes librerías “caseras”. Illumina® ha propuesto recientemente una alternativa comercial a los paneles de genes: el kit de captura TruSight One Sequencing Panel para 4813 genes implicados en enfermedades humanas hereditarias. El uso de este kit permitiría reducir los costes del análisis y uniformar el diagnóstico entre los varios laboratorios que trabajan en el diagnóstico de las enfermedades raras. Pero, como demuestra nuestra tesis, la gran ventaja de la WES respecto a la NGS de paneles de genes, es que permite acoplar la investigación de base con el diagnóstico. La WES, secuenciando indiscriminadamente todo el exoma, permite el descubrimiento de nuevos genes cuyas mutaciones puedan ser causa de la enfermedad y además, cuando en un futuro se habrán secuenciado suficientes exomas de pacientes FA, permitirá eventualmente proceder a la búsqueda de variantes génicas que, aun sin ser la causa de la enfermedad, podrían modular su severidad.

OBJETIVOS

“If you’re committed enough, you can make any story work. I once told a woman I was Kevin Costner, and it worked because I believed it”.

Saul Goodman

III. OBJETIVOS

Esta tesis doctoral es el producto directo de la línea de investigación que el Grupo de Inestabilidad Genómica y reparación del DNA del Departamento de Genética y Microbiología de la Universidad Autónoma de Barcelona viene desarrollando desde la creación en 1998 del Registro Español de Pacientes Fanconi. Desde entonces la “Red Nacional de Investigación Cooperativa sobre FA” (<http://www.redfanconi.org>) ha recolectado casi la totalidad de los pacientes españoles de FA. Dentro de esta población quedaban un 8% de pacientes cuyos grupos de complementación no se pudieron encontrar y que por lo tanto podían tener mutaciones en genes desconocidos, y también había un 30% de pacientes parcialmente subtipados y caracterizados. Además, casi cada mes, nuestro laboratorio confirma el diagnóstico de un nuevo paciente de FA. Por lo tanto decidimos de aplicar las tecnologías de NGS, y en particular la WES, como alternativa a los métodos clásicos de subtipaje y de caracterización de las mutaciones en una cohorte de 49 pacientes FA. Nuestro objetivo fue ensayar la WES en las condiciones más estándar posibles utilizando los servicios de empresas externas y kit de captura comerciales convencionales para averiguar si era posible caracterizar los pacientes FA de nuestra cohorte de la forma más rápida, sencilla y económicamente sostenible.

También quisimos utilizar la WES como herramienta para el descubrimiento de nuevos genes y en particular en dos pacientes FA que ya se habían estudiado de forma muy completa pero sin ningún resultado positivo que identificara el gen mutado. Hasta entonces todos los intentos de caracterizar los genes mutados en estos dos pacientes habían fracasado. Una vez descartados todos los genes ya conocidos, se utilizaron, sin ningún resultado positivo, arrays de SNPs e Hibridación Genómica Comparada (CGH) para encontrar en el genoma regiones de homocigosis o regiones delecionadas donde se pudieran esconder los genes responsables de la enfermedad. También se intentó una aproximación por genes

candidatos y se secuenciaron inútilmente muchos genes ya conocidos por tener un papel en la resistencia a los ICL en varios organismos.

En el caso de encontrar un nuevo gen FA, otro objetivo de esta tesis sería verificar su papel en la predisposición al cáncer de mama y ovario familiar.

En concreto nuestros objetivos fueron:

1. La caracterización mediante WES (sistemas SOLiD4TM e IlluminaTM) de un panel de 49 pacientes de FA no caracterizados o parcialmente caracterizados.
2. Conseguir una prueba de principio de la utilidad de la WES en el subtipaje y la caracterización molecular de los pacientes FA con los kits de captura comerciales más utilizados.
3. Encontrar genes nuevos cuyas alteraciones causan FA en pacientes donde ya se habían descartado todos los genes FA hasta ahora conocidos.
4. Determinar el papel molecular de las proteínas codificadas por los nuevos genes identificados en la ruta FA/BRCA de reparación del DNA.
5. Determinar la implicación de los nuevos genes identificados en la susceptibilidad familiar al cáncer de mama y ovario.

MATERIALES Y MÉTODOS

“That's another fine mess you've gotten me into.”

Oliver Hardy

IV. MATERIALES Y MÉTODOS

IV.1. Pacientes y controles

El diagnóstico de la FA fue confirmado para los 49 pacientes mediante el test de fragilidad cromosómica inducida por DEB (Auerbach, 2003, Castella et al., 2011a). Los pacientes fueron clasificados como mosaicos cuando tenían un porcentaje de células aberrantes inferior a 50-60%. Se consiguió el consentimiento informado de todos los pacientes incluidos en este estudio en conformidad con la Declaración de Helsinki. Esta investigación fue aprobada por el Comité Ético de la Universidad Autónoma de Barcelona para la Investigación Humana.

El paciente FA104 nació en España en el 2000 de padres no emparentados y fue diagnosticado en el periodo neonatal por un síndrome malformativo sugerente FA. El paciente presentaba la ausencia de los dos pulgares, microsomía, atresia del esófago, ano anterior, hipotiroidismo y oídos dismórficos. El paciente no presentaba anomalías dermatológicas tales como hiperpigmentación de la piel, fotosensibilidad o epidermis atrófica. El inicio de la enfermedad hematológica fue a la edad de 2 años, con una anemia aplásica que afectó a los tres linajes sanguíneos reconducible a una insuficiencia de la médula ósea. El paciente se sometió a un trasplante de médula ósea a la edad de 4 años y murió debido a un shock hemorrágico después del trasplante. La prueba de fragilidad cromosómica fue claramente positiva lo que confirmaba el diagnóstico de FA. La LCL derivada de este paciente era hipersensible a MMC y Melfalán, y mostraba una monoubiquitinación normal de FANCD2 y de *foci* normales de RAD51. Esto era sugerente de un posible defecto en la vía FA/BRCA por debajo de la monoubiquitinación de FANCD2 que pero no implicaba alteraciones de BRCA2 o de la recombinación homóloga (Casado et al., 2007).

El paciente 1333 nació en Alemania en 2002 y fue diagnosticado de FA, a la edad de 5 años. Los síntomas clínicos incluían: bajo peso al nacer y retraso del crecimiento postnatal, baja estatura, microcefalia pronunciada, manchas “café con leche”, defecto del *ostium primum* (agujero de comunicación entre ambas las aurículas del corazón fetal), hipoplasia del tracto biliar con fibrosis del hígado e insuficiencia de la médula ósea. De forma parecida al paciente FA104, el 1333 no presentaba ni fotosensibilidad de la piel ni epidermis atrófica y la prueba de fragilidad cromosómica era claramente positiva. Los controles consistieron en 200 representantes sanos de la población española, reclutados principalmente desde el Centro de Investigación de la menopausia en el Instituto Palacios (Madrid, España) y del Colegio de Abogados (Madrid, España). Los detalles de esta serie de controles han sido publicados anteriormente (Milne et al., 2006).

Para investigar el posible papel de *ERCC4* en la susceptibilidad al cáncer de mama y ovario, se analizaron 1573 individuos provenientes de familias con un historial de cáncer de mama y ovario que ya habían dado negativo para mutaciones en *BRCA1/FANCS* y *BRCA2/FANCD1*. Los criterios de inclusión en la población fueron: tener por los menos tres parientes de primer grado con cáncer de mama u ovario o dos mujeres parientes de primer grado con cáncer de mama (por lo menos una diagnosticada antes de los 50 años) o tener entre los parientes de primer grado un cáncer de mama más un caso de cáncer de ovario o un caso de cáncer de mama bilateral o un caso de cáncer de mama en un varón. Los controles fueron 854 personas sin historial de cáncer personal ni familiar ya precedentemente descritos (Bonache et al., 2013, Fernandez-Rodriguez et al., 2012, Osorio et al., 2012, Romero et al., 2011).

IV.2. Ensayo de fragilidad cromosómica

El ensayo de fragilidad cromosómica se realizó básicamente tal y como se propone en (Auerbach, 1994). Para cada línea celular se establecieron 3 cultivos en las condiciones adecuadas para el tipo celular. Se trataron dos de los cultivos a las dosis indicadas de DEB y uno se utilizó como control de daño espontáneo. A las 48-72h del tratamiento se procedió a la preparación de metafases, habiendo añadido 2 horas antes Colcemid a una concentración

final de 0,1 µg/ml. Las preparaciones de las metafases se realizaron según el método habitual: el botón de células se resuspendió en solución hipotónica (0,075M KCl) que se dejó actuar durante 25 min a 37°C. Se fijaron las células mediante 3 lavados con una solución de metanol y ácido acético, en una proporción 4:1. El material obtenido fue goteado sobre portaobjetos que, posteriormente, se tiñeron con Giemsa. Para el recuento de las aberraciones, se contaron 50 metafases de cada cultivo con el sistema Metafer Slide Scanning Platform de Metasystems.

IV.3. WES y análisis de las mutaciones en pacientes FA.

El DNA genómico analizado fue aislado mediante un método estándar, de sangre periférica (*peripheral blood*; PB), fibroblastos primarios o LCLs inmortalizadas. En el caso de PB se procedió a una lisis inicial de los eritrocitos mediante una solución hipotónica y a la recuperación de un botón de linfocitos que posteriormente se sometió a lisis celular con un tampón específico incubando a 37° durante 16h con Proteinasa K y RNAsa para la digestión de las proteínas y del RNA. El DNA se recuperó mediante la adición de NaCl y fue separado de las proteínas y demás restos celulares mediante la adición de cloroformo. Finalmente, el ADN se precipitó con etanol y se resuspendió en ddH₂O.

Para la secuenciación del exoma mediante sistema SolidTM 4 se utilizaron los servicios ofrecidos por la empresa “Sistemas Genómicos” de Valencia (España). 3 µg de DNA de los pacientes FA104 y FA287 se utilizaron para generar una librería de fragmentos según el método de preparación SOLiDTM 4. El enriquecimiento de las secuencias exónicas se realizó utilizando el sistema en fase líquida Human All Exon Target Enrichment System for 38 Mb (Agilent). Los fragmentos se amplificaron por emPCR siguiendo las instrucciones del fabricante y se secuenciaron según los estándares de SOLiDTM 4. 107,8 millones de micro esferas se depositaron en un cuarto de un portaobjeto de secuenciación y se produjeron un total de 208.265.974 lecturas *paired-end* asimétricas. Las combinaciones de colores (*color-space*) fueron alineadas al genoma humano de referencia (NCBI36/hg18) utilizando Bioscope v1.2 (<http://solidsoftwaretools.com>) con la configuración predeterminada. Las variantes de único nucleótido (SNV) fueron identificadas con Samtools (Li et al., 2009)

utilizando los siguientes criterios de filtrado: calidad de SNV de 20 (Phred-like score), calidad de genotipo de 30 (Phred-like value) y una cobertura mínima de 9x. Inserciones y deleciones (InDel) fueron identificadas con la herramienta Small Indel que es parte de la suite de análisis Bioscope v1.2. Se aplicó un único filtro de al menos 9 lecturas no redundantes para las InDel. Las variantes identificadas se anotaron y se clasificaron según su posición o efecto sobre de los transcritos mediante la Application Programming Interfaces de Ensembl (Rios et al., 2010) para compararlas con las en la base de datos Ensembl (versión 59) que incluye datos de dbSNP, del proyecto 1000Genome y del proyecto HapMap. Las nuevas variantes que resultaban en un cambio de la transcripción no-sinónimo, o en un frameshift o en un cambio no frameshift y alteraciones ubicadas dentro de los sitios de splicing se consideraron para identificar nuevos genes candidatos (Meindl et al., 2010).

Para la secuenciación del exoma mediante Illumina las librerías de los pacientes se crearon utilizando el kit Agilent SureSelect All Exon V5 a partir de 1 µg de DNA de los pacientes, siguiendo las instrucciones del fabricante. Las librerías se secuenciaron con un Illumina HiSeq y los datos crudos se alinearon al genoma de referencia (human, hg19, GRCh37) con BWA (Li and Durbin, 2009). Solamente las secuencias únicas se tomaron en consideraciones para los análisis sucesivos. La secuencias se analizaron con el Genome Analysis Toolkit (GATK) (McKenna et al., 2010). Las SNV y las InDel se identificaron con el GATK UnifiedGenotyper y los efectos de las variantes se determinaron con snpEff.

Todas las variantes de interés detectadas por WES, fueron posteriormente confirmadas mediante secuenciación clásica de Sanger del DNA mediante el kit BigDye Terminator Cycle Sequencing V3.1 Ready Reaction (Applied Biosystems) y analizadas en un secuenciador ABI PRISM 3100.

IV.4. MLPA

La técnica de la MLPA para la detección de deleciones se realizó básicamente como está descrito en (Schouten et al., 2002). Para ello, se empleó el kit de detección de MRC-Holland BV, Amsterdam, para el gen *FANCA* (detalles disponibles en <https://www.mrc-holland.com>). Brevemente, el DNA es desnaturalizado 5min a 98°C, se añaden 3µl de mezcla

de las sondas específicas para cada exón y se deja incubar 16h a 60°C. Pasado este tiempo se añade la ligasa y se incuba a 54°C, 15min. Seguidamente, 10µl de esta reacción se transfieren a un tubo que contiene la mezcla de amplificación para la PCR (tampón de PCR, dNTPs, polimerasa Taq y los cebadores universales, de los cuales, uno está marcado con carboxifluoresceína. La amplificación se lleva a cabo en 33 ciclos (30seg a 95°C, 30seg a 60°C y 60seg a 72°C). Los fragmentos son detectados mediante un secuenciador ABI310 (Applied Biosystems) y los datos se analizan con el software COFFALYSER.NET (mrc-holland.com).

IV.5. Líneas celulares

Todas las líneas celulares utilizada en nuestro estudio se reportan en la Tabla III. Las LCLs humanas transformadas con virus de Epstein-Bar (Pattengale et al., 1973, Frisan et al., 2001) se cultivaron en Roswell Park Memorial Institute medium (RPMI, JRH Biosciences) suplementado con 15% de Suero Fetal Bovino (FCS) y con el antibiótico Plasmocin (Invivogen) a la concentración final de 2,5 µg/ml. Los fibroblastos humanos inmortalizados se cultivaron en Dulbecco's Modified Eagle's Medium (DMEM, (JRH Biosciences)) suplementado con 10% de FCS y Plasmocin (2,5 µg/ml). Los fibroblastos primarios humanos se cultivaron en DMEM con 15% FCS y plasmocin (2.5 µg/ml). Las líneas GM02344, GM08437 (XP2YO^{SV40}), GM16633 (PD20), GM16634 (PD20-3.15), GM04313 (XP2YO^{PF}) se adquirieron de Coriell Cell Repositories. Las líneas de fibroblastos embrionarios de ratón inmortalizados (MEF) *wild type* (wt) y *Ercc4^{-/-}* fueron un regalo de la Dr. María A. Blasco del Centro Nacional de Investigaciones Oncológicas (CNIO) de Madrid y se cultivaron en Dulbecco's Modified Eagle's Medium (DMEM, (JRH Biosciences)) suplementado con 10% de FCS y Plasmocin (2,5 µg/ml concentración final). Todas las líneas celulares se conservan en N₂ líquido en el Biobanco de síndromes de la reparación del DNA de nuestro grupo.

Tabla III: Líneas celulares utilizadas en el estudio:

	Código	Tipología		Código	Tipología
	Líneas celulares linfoblastoides inmortalizadas con EBV	FA139		wt	Líneas celulares fibroblástoides inmortalizadas con SV40
HSC93		wt	PD20	FA-D2	
FA104		FA-Q	PD20-3.15	FA-D2 + FANCD2 wt	
1333		FA-Q	1333 ^{SV40}	FA-Q	
FA104R		FA-Q revertida			
1333R		FA-Q revertida	Líneas de fibroblastos primarios	Código	Tipología
HSC536		FA-C		XP2YO ^{PF}	XP-F
EUFA1354		FA-P		XP42RO	XP-F
475/3		FA-P		XP51RO	XFE
GM02344		XP-A		FA90	FA-A
HSC230NT		FA-B	1333 ^{PF}	FA-Q	
EUFA867		FA-M	Fibroblastos embrionarios de ratón inmortalizados	Código	Tipología
EUFA579		FA-D1		MEF wt	wt
EUFA696		FA-J		MEF <i>Ercc4</i> ^{-/-}	<i>Ercc4</i> ^{-/-}

IV.6. Análisis del bloqueo del ciclo celular en G2 y citometría de flujo

Se sembraron 1×10^6 células en las condiciones adecuadas según el tipo celular y se les añadió MMC a la concentración final de 50 nM. Al mismo tiempo se estableció también un cultivo no tratado para la evaluación del ciclo celular en condiciones basales. Pasadas 72h se realizó el recuento de las células para evaluar la viabilidad y verificar que las células no tratadas hubiesen proliferado durante este tiempo. Después de eliminar del medio de cultivo y lavar con PBS, las células se recolectaron y se fijaron en etanol 70% frío durante 30 min a 4°C. El ciclo celular se evaluó mediante citometría de flujo, midiendo la cantidad total de DNA en cada célula: la muestra se preparó eliminando el etanol y resuspendiendo 1×10^6 células en 1 ml de una solución de tinción (Ioduro de Propidio [40µg/ml], RNAsa A [0,1mg/ml] en PBS). Después de una incubación de 30 min a 37°C, las células se analizaron por citometría de flujo en un citómetro modelo FACScalibur (Becton Dickinson). Los porcentajes de células que se encontraban en cada una de las fases del ciclo celular se determinaron mediante el programa informático FlowJo (Tree Star Inc.).

IV.7. Clonaje específico de alelos

Los exones 8 y 11 de *ERCC4* fueron amplificados por PCR a partir del DNA genómico de la línea FA104R. Los productos de PCR se purificaron (NucleoSpin® Gel and PCR Clean-up, MACHEREY-NAGEL) y se clonaron con el Topo TA cloning kit (Invitrogen) siguiendo las instrucciones del fabricante. Los plásmidos con los productos de PCR insertados se transfectaron en células competentes DH5 α . Las células fueron sembradas en placas de LB con ampicilina y se picaron 100 colonias para las preparaciones de los plásmidos. El DNA purificado (NucleoSpin® Plasmid QuickPure Kit (Macherey-Nagel)) se secuenció para verificar la presencia de mutaciones producidas por el proceso de selección que fueran capaces de revertir el efecto de las mutaciones presentes en los paciente. En la Tabla IV se describen las secuencias de todos los cebadores utilizados para la secuenciación a nivel genómico o a nivel del cDNA del gen *ERCC4*.

IV.8. Ensayo de sensibilidad a los ICL y UVC en células adherentes.

Se sembraron 5×10^4 células en placas de Petri de 60 mm y se dejaron adherir por 16-18 horas, entonces se cambió el medio de cultivo por medio con MMC (Sigma) a las diferentes concentraciones analizadas. Para cada punto se utilizaron dos placas para minimizar la variación experimental. Las células se dejaron crecer hasta que el control no tratado había realizado por lo menos tres (fibroblastos primarios) o seis (MEF y fibroblastos humanos inmortalizados) doblajes poblacionales. Entonces las células se lavaron con PBS y se recogieron por tratamiento con tripsina. El número de células en la suspensión celular se contó en un contador de células automático Z2™ COULTER COUNTER (Beckman Coulter). Se llevaron a cabo por lo menos tres experimentos independientes y los resultados se expresaron como porcentaje de células tratadas viables en relación con las de los controles no tratados. Para el ensayo de inhibición de la proliferación por UVC se procedió a la siembra de las células como descrito arriba y se dejó que se pegaran al substrato por 16-18 horas. Entonces las células se lavaron con PBS y se irradiaron con UVC por medio de una lámpara Philips 15 W UVC G15-T8 a las dosis indicadas. Las células se dejaron crecer y

fueron procesadas y contadas como en el caso del ensayo de inhibición de la proliferación por ICL.

Tabla IV: Cebadores para la secuenciación de *ERCC4* y para el clonaje específico de alelos

Exón:	Cebadores:
Exon 1 for	ACTCGGCTCTCTCGGTTGAGTT
Exon 1 rev	CGCTCAGGAGGCCCTCAACA
Exon 2 for	AACTGCCCTGTATTAATAGCCTACTAA
Exon 2 rev	GTACAAATTTACATACTAATAATAAGATTTCA
Exon 3 for	AATGTGATGAATGAATGGCAATTACCTAC
Exon 3 rev seq	TTACATCAAAGTGCAACTAAGTTACAGTAG
Exon 3 rev	GTTTGTATAAACTGACTAGGTTCTATCA
Exon 4 for	CATAGCTGCTGAAACTCTAGAAAATTGTTGAAA
Exon 4 rev	AGTCAGAGTGATGCTTATATGCCAATCCAC
Exon 5 for	TACACAGGAAATAATCCTTTTGAAAGTATG
Exon 5 rev	CAGTATAACATATAGTTGAATATAGCACTT
Exon 6 for	ACAGGATGACAGCCAGTTACGTATGTAG
Exon 6 for seq	TAGGTCATGTGACCATCAGAGACTGTT
Exon 6 rev	ACACATTTAAAGACTTAACCCCAACAAG
Exon 7 for	TATGTAAGTATGCTCGTGTATCTGTTGTT
Exon 7 for seq	GTTTTAAAAGCCTTTGGAAGACTTTATGG
Exon 7 rev	CACTAGGATCTCAGTGTTCAATTTGCCAT
Exon 8 for	TAATACCAAAGGGTAAGATGTCTTCCCTT
Exon 8 for seq	GGTGAAGGAATAAGGGGGCAC
Exon 8 rev	CATAAACATAAGCAGCATCGTAACGGATAT
Exon 9 for	AATATTTGTTATTTGAGCGCTCTAGGTTGC
Exon 9 rev	CCTGAGCAGGACTATCTGATATTCC
Exon 10 for	ACTGCTATCATCATGTAGATCATTTCATAC
Exon 10 rev	TACAGACCAAGCCTTGGCAGAGTAACT
Exon 11a for	AGAGTTAACAACAGAAACATCGATTTTTAG
Exon 11a rev	CTCAAACAACCTCCGCCGTTGCATGA
Exon 11b for	CTTGTTTCAGGAGATCTCCAGCAATG
Exon 11b rev	TATGATGTCTGGCAAGGAGCCGCT

IV.9. Ensayo de la sensibilidad a los ICL y UVC en LCLs

Para la determinación de la sensibilidad a los ICL de las LCLs, se sembraron en placas de seis pocillos 5×10^5 células por pocillo en 2 ml de medio completo. Se añadió MMC a las concentraciones indicadas y se dejaron crecer las células hasta que el control hubiera

realizado por lo menos tres doblajes poblacionales. Las células se recolectaron y se contaron con el contador de células y partículas automático Z2™ COULTER COUNTER (Beckman Coulter). Se llevaron a cabo por lo menos tres experimentos independientes y los resultados se expresaron como porcentaje de células tratadas viables en relación a los controles no tratados. Para la determinación de la sensibilidad a la radiación UVC de las LCLs se utilizó el siguiente método: se recolectaron 1.5×10^6 de células y se lavaron dos veces con PBS. El botón de células se resuspendió en 2 ml de PBS al 3% de FCS y la suspensión fue distribuida uniformemente en una placa de Petri de 10 cm de diámetro. Las células se irradiaron a las dosis indicadas y por cada punto se utilizaron dos placas para minimizar la variación experimental. Se añadieron 3 ml de medio completo y se recuperaron las células por centrifugación. Entonces, las células se resuspendieron en 5 ml de RPMI completo y se dejaron recuperar por 24 h en condiciones de cultivo. Después de las 24 horas, las células fueron procesadas para evaluar el apoptosis con el Annexin-V-FLUOS Staining Kit (Roche) según las instrucciones del fabricante. Se analizaron 20,000 células por citometría de flujo (FACS Canto, BD Biosciences) y los resultados fueron analizados con el programa informático Flowjo (Tree Star, Inc.). Los resultados se expresaron como el porcentaje de células vivas (negativas por ambas las señales de la anexina y del ioduro de propidio) respecto al total de las células y en relación al control no tratado.

IV.10. Producción de las líneas revertidas FA104R and 1333R.

Las LCLs FA104 y 1333 se cultivaron en presencia constante de MMC a la concentración de 15 nM (FA104) o de 40 nM (1333) hasta que las células empezaron a proliferar de forma estable indicando que los cultivos habían perdido la sensibilidad a los ICL. La resistencia a los ICL de las líneas revertidas se confirmó con ensayos de inhibición de la proliferación celular en presencia de MMC y de parada en la fase G2M del ciclo celular. Las LCLs que habían vuelto a ser resistentes a la MMC se renombraron FA104R y 1333R.

IV.11. Mutagénesis sitio dirigida

Para la mutagénesis sitio-dirigida, los cebadores mutagénicos se diseñaron utilizando el software de Agilent Technologies (Tabla V). Los cebadores mutagénicos deben contar con alrededor de 10 a 15 bases de secuencia correcta de ambos los lados de la base que se quiere cambiar, de modo que se puedan hibridar correctamente en la siguiente reacción de síntesis. Para introducir las variantes, se utilizó como molde el plásmido lentiviral pWPXL con el cDNA wt de *ERCC4* insertado junto a una secuencia HA (derivada del gene para la glicoproteína hemoaglutinina del virus de la influenza humana) como epítipo en el extremo COOH y se utilizó el QuickChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies). Después de la amplificación por PCR, se añadió la enzima de restricción DpnI durante 1 hora a 37° con el fin de digerir el DNA molde de doble cadena. DpnI solo corta el DNA en sitios metilados, por lo tanto elimina el molde, que es metilado por las bacterias, pero no el producto de la PCR que no está metilado. Después de la digestión se transformaron las células XL10-Gold ultra competentes (Agilent Technologies) y se pusieron a crecer a 37°C hasta al día siguiente en placas de LB-agar con la selección adecuada para el plásmido (Figura 13). Se picaron las colonias y el día siguiente se preparó el DNA plasmídico mediante el kit NucleoSpin® Plasmid QuickPure (Macherey-Nagel). Se controló por secuenciación todo el cDNA de *ERCC4* para verificar que el proceso de mutagénesis no hubiese incluido algún otro cambio no deseado en la secuencia del gen. El cDNA de los varios mutantes de XPF se clonó sucesivamente en el vector retroviral pBABE IRES *EGFP*.

Para generar el cDNA con la duplicación de 28 pares de bases (bp) se procedió a amplificar el vector lentiviral pWPXL-XPF con los siguientes cebadores:

5'CGATAATCCGGACTTACTTCACTTCCCCAGACTACGG 3'

5'GAAGTACTCCGGATGCAGCTCTCGGGC 3'

En una segunda PCR se utilizaron los siguientes cebadores:

5'GAGCTGCATCCGGAGTACTTCAAG3'

5'GTTA ACTCCGGAGTCTGGGGAAGTGAAGTGTAAGAAG3'.

Se cortó el producto de las dos amplificaciones con la enzima BmpE1 para generar las extremidades cohesivas y, después de la reacción de ligación, se transformaron y se analizaron los plásmidos resultantes por secuenciación. El plásmido con el inserto correcto se utilizó como molde para añadir mediante PCR el epítipo HA al extremo C-terminal de la proteína XPF mutante. Para esta PCR se utilizaron los siguientes cebadores:

5'TAATATACGCGTATGGAGTCAGGGCAGCCGGCTCGACGGATTGCCATGGCGCC
GCTGCTGGAGTAC3'

5'AATCTAACTAGTTCAAGCGTAATCTGGAACATCGTATGGGTAAACAACCTCCGC
CGTTGCATGAGG3'

El producto de la PCR y el vector se digirieron con las enzimas de restricción MluI y SpeI y se ligaron generando, por fin, el pWPXL-XPF-28bp-HA. Todos los plásmidos se verificaron por secuenciación.

Tabla V. Secuencias de los cebadores utilizados para generar las mutaciones puntuales encontradas en *ERCC4* en los pacientes FA y en la población de cáncer de mama hereditario.

Mutación	Cebadores
p.Arg689Ser	GCATAGTTGTGGATATGAGTGAATTCGAAGTGAGCTTCC GGAAGCTCACTTCGAAATTCACATATCCACAACATGC
p.Leu230Pro	ATGCTAGCTATACAGACTGCTATACCGGACATTTTAAATGCATGTCTAAAG CTTTAGACATGCATTTAAAATGTCCGGTATAGCAGTCTGTATAGCTAGCAT
p.Arg150Cys	CAAGAAGCATTATCTTGTGCCTCTTTCGCCAGAAAA TTTTCTGGCGAAAGAGGCACAAGATGAATGCTTCTTG
p.Ser786Phe	GTCGGGGCATTGACATCGAACCCGTGACTTTAG CTAAAGTCACGGGTTCGATGTCAATGCCCGAC

IV.12. Vectores y producción de partículas virales

A lo largo de nuestro estudio se utilizaron los siguientes vectores retrovirales: LFAPEG-wt*FANCA*, pBABE-IRES-*EGFP* y pBABE-IRES-*EGFP* donde se clonaron los varios mutantes de *ERCC4* generados por mutagénesis sitio dirigida como se ha explicado anteriormente. También se emplearon los siguientes vectores lentivirales: pWPXL-*EGFP* y

pWPXL donde se clonaron los varios mutantes de *ERCC4*. Para la producción de partículas retrovirales aptas para infectar células derivadas de ratón se utilizó la línea celular empaquetadora Platinum E (Cell Biolabs): 5×10^6 células platinum E se sembraron en placas de 10 cm de diámetro en DMEM (JRH Biosciences) con 10% de FCS y plasmocin 2,5 $\mu\text{g}/\text{ml}$ un día antes de la transfección. El día siguiente el medio de cultivo se cambió por medio fresco y 2 h después cada placa se transfectó con 30 μg del plásmido de interés utilizando Lipofectamine 2000 (Life Technologies). Se cambió el medio después de 8 h y se recogió el medio que llevaba las partículas retrovirales después de otras 48 h. El sobrenadante se limpió centrifugándolo a baja velocidad (200 Gs) por 10 min y se filtró con filtros de policarbonato de 0,45 μm (Millipore). Este sobrenadante se utilizó para infectar las células de interés sembradas 24 h antes de la infección en placas de 6 pocillos (1×10^5 /pocillo). Para la producción de las partículas lentivirales, se procedió como en (Almarza et al., 2007). Para la infecciones de las células se procedió como en el caso de las infecciones con las partículas retrovirales utilizando 20 μl de virus concentrando en 2 ml totales de medio de cultivo. En el caso de las infecciones de las LCLs las placas multi-pocillos estaban recubiertas de fibronectina (2 $\mu\text{g}/\text{cm}^2$, Takara Shuzo) (5×10^5 LCLs/pocillo). Se hicieron tres rondas de infección cada 24 h para conseguir niveles de transducción adecuados. Los fibroblastos humanos inmortalizados se transdujeron una sola vez en placas de 6 pocillos (1×10^5 células/pocillo) con 20 μl de sobrenadante viral en 2 ml totales de medio cultivo.

IV.13. Selección de células por fluorescencia

Después de la selección con puromicina, las seis líneas MEF *Ercc4*^{-/-} transducidas con el vector vacío pBABE-IRES-EGFP y con el mismo vector donde se habían clonado los mutantes p.Arg689Ser, p.Leu230Pro, R150C y S786F y el cDNA normal de *ERCC4*, se seleccionaron por citometría de flujo (*Fluorescence-Activated Cell Sorting*; FACS) para conseguir una población homogénea que expresase la GFP y por lo tanto también las distintas formas de XPF. El *Internal Ribosome Entry Site* (IRES), es una secuencia de nucleótidos que permite la iniciación de la traducción en el medio de una secuencia de un RNA mensajero. Estas secuencias fueron descubiertas por primera vez en el virus de la

encefalomiocarditis y en el poliovirus, ambos virus de RNA (Jang et al., 1988, Pelletier and Sonenberg, 1988). Cuando un segmento IRES se encuentra entre dos marcos de lectura en una misma molécula de mRNA (mRNA bicistrónico), se puede producir la traducción de la proteína codificada después del IRES independientemente de la estructura 5'-Cap de la molécula de mRNA. En el caso de nuestros constructos la secuencia IRES se encuentra entre la ORF de XPF y la ORF de la GFP y ambas proteínas se producen por separado en la célula aunque estén codificadas por una única molécula de mRNA. Las líneas celulares se seleccionaron utilizando un citómetro FACS Aria II (BD Biosciences), equipado con láser azul de 488 nm para detectar la fluorescencia de GFP y se purificó una población compuesta por los menos de un 98% de células que expresaban la EGFP.

IV.14. Ensayo de síntesis de DNA no programada

La síntesis de DNA no programada (*Unscheduled DNA Synthesis*; UDS) es la replicación del DNA durante los procesos de reparación de daños al DNA y como tal, es distinta de la replicación del DNA de la fase S del ciclo celular eucariota. La valoración de la UDS después de irradiación de las células con UVC es el ensayo que se utiliza en el diagnóstico de la XP para evaluar la eficiencia de NER (Stefanini et al., 1980, Lehmann et al., 2011). En los métodos más recientes se utilizan marcadores no radioactivos como la 5-ethynyl-deoxyuridine, un análogo de la desoxiuridina (Kleijer et al., 2006, Limsirichaikul et al., 2009). Para la medida de la UDS a lo largo de esta tesis doctoral, las células se cultivaron por tres días en presencia de micro esferas de poliestireno cuya superficie es recubierta de grupos carboxilos (Polysciences) de 0,75 μm de diámetro para las células donde se quería evaluar la eficiencia de la reparación por NER y de 2 μm para las células control. Se recolectaron las células, se mezclaron en proporción de 4: 1 (células analizadas: controles) y se pusieron otra vez en cultivo sobre de portaobjetos para microscopia. Este proceso permite de tener un control interno, que se puede reconocer fácilmente por el tamaño de las esferas dentro del citoplasma de las células, que consiente de eliminar las variaciones entre las diferentes preparaciones de los portaobjetos para la microscopia (Vermeulen et al., 2000, Jaspers and Bootsma, 1982). Estos cultivos celulares “mixtos” se irradiaron con 16 J/m² de

UVC como en el caso de los ensayos de inhibición de proliferación y se incubaron por 3 h (fibroblastos primarios) o 1 h (MEF) en medio de cultivo en presencia de 5'-etnil-deoxiuridine 20 μM y 5-fluorodeoxiuridina 1 μM (ClickIt kit, Invitrogen). Después de una incubación de 15 min en medio que contiene timidina 10 μM , las células se fijaron en 3,7% de formaldehído en PBS y se permeabilizaron con PBS 0,5% Triton-X100 antes de añadir la azida acoplada al fluorocromo Alexa Fluor 594 siguiendo las instrucciones del Kit (ClickIt kit, Invitrogen). Se tomaron fotos y utilizando el software ImageJ (NCBI) se midió la intensidad de la fluorescencia de 20-40 núcleos elegidos al azar que no estuviesen ni en mitosis ni en fase S.

IV.15. Irradiación local con UVC y reparación del daño al DNA

Se cultivaron las células inmortalizadas XP2YO^{SV40} que expresaban las diferentes formas de XPF sobre cubreobjetos para microscopia y se irradiaron con UVC (254 nm) a 150 J/m² por medio de una lámpara Philips 15 W UVC G15-T8 a través de un filtro de policarbonato con poros de 5 μm (Millipore) como ya descrito en (Bogliolo et al., 2007). Para la detección de los fotoproductos pirimidina (6-4) pirimidona (6-4PP) se utilizó un anticuerpo monoclonal de ratón (Cosmo Bio) diluido 1:400 en solución de bloqueo mientras que para detectar XPF se utilizó el anticuerpo policlonal de conejo (ab9110, Abcam) contra el epítipo HA a la dilución de 1:3000. Las muestras se incubaron a 4°C por 18 h y después de tres lavados de 10 min con solución de bloqueo, se incubaron con los siguientes anticuerpos secundarios: *goat anti-mouse* conjugado con Cy3 (Jackson ImmunoResearch) 1:1,000 y *goat anti-rabbit* Alexa Fluor 488 (Invitrogen) 1:800 por 1 h a temperatura ambiente. Después de 2 lavados en PBS 0,01% Tween de 10 min y un lavado en ddH₂O se dejaron secar los cubreobjetos protegiéndolos de la luz directa y se montaron en los portaobjetos con Vectashield Mounting Medium con 1.5 $\mu\text{g/ml}$ of DAPI (Vector Laboratories). Las preparaciones fueron analizadas con un microscopio confocal Zeiss LSM 510. Se contaron alrededor de 100 células por cada punto y se realizaron dos experimentos independientes. Para el análisis de la reparación de los daños al DNA en LCLs, las células se

resuspendieron en RPMI 50% FCS y se sembraron en cubreobjetos recubiertos de Poli-L-lisina (Sigma). Se dejaron adherir por 1 h a 37°C y después de eliminar el medio, se lavaron con PBS. Para la irradiación (60 J/m^2) y la inmunofluorescencia (IF) se procedió como en el caso de los fibroblastos humanos inmortalizados. Los anticuerpos utilizados fueron un monoclonal de ratón contra los dímeros de ciclobutano de pirimidinas (*Cyclobutane pyrimidine Dimers*; CPD) diluido 1:100 (Kamiya Biomedical Company) y un policlonal de ratón contra XPF a la dilución de 1:50 (ab76948, Abcam).

IV.16. Western Blot, co-IP y fraccionamiento celular

Para la preparación de extractos totales de proteínas celulares se lisaron las células en solución RIPA (Millipore) complementada con un cóctel de inhibidores de proteasas (CompleteMini-EDTA free, Roche) e inhibidores de fosfatasas (PhosSTOP, Roche). Para la degradación de la cromatina, se añadió a la muestra 25U/ml de Benzona (Novagen) y se incubó a temperatura ambiente durante 10 min. Los extractos se centrifugaron a 16000 RPM en una centrifuga de poyata y la concentración de proteínas totales del sobrenadante se determinó mediante el sistema Bio-Rad Protein Assay (Biorad), según las instrucciones del fabricante. Se cargaron 50 µg de proteínas totales en un gel desnaturizante SDS-PAGE a la concentración de acrilamida apropiada para las proteínas que se querían analizar y se procedió con un método de Western Blot (WB) estándar. Se transfirieron las proteínas por 1 h a 20 V a una membrana de nitrocelulosa (Biorad) utilizando el sistema Transblot Semi-Dry de Biorad y posteriormente, se realizó una incubación de la membrana con solución de bloqueo (5% de leche desnatada en polvo en TBS+Tween) por 1 h a temperatura ambiente. La inmunodetección se realizó con los anticuerpos primarios diluidos en solución de bloqueo durante 16 h a 4°C. Los anticuerpos primarios utilizados fueron: anti-HA (ab9110, Abcam) 1:1000, anti-Vinculina (ab18058, Abcam) 1:3000 y anti-XPF (ab76948, Abcam) 1:1000. Posteriormente se incubó la membrana con un anticuerpo secundario adecuado acoplado a peroxidasa (Abcam), durante 2 h a 4°C y se detectó la señal después de incubar la membrana con el Pierce ECL Western Blotting Substrate (Pierce). Para las co-IP se lisaron 10×10^6 células en tampón de lisis (50 mM Tris HCL (pH 7.4), 150 mM NaCl y 1% Triton X-100,

complementado con inhibidores de proteasas (EDTA-free Protease Inhibitor Cocktail tablets, Roche) y fosfatasas (PhosSTOP Phosphatase Inhibitor Cocktail tablets, Roche). Se inmunoprecipitaron las proteínas por 2 h a 4°C con anticuerpos contra SLX4 y ERCC1 y sucesivamente se incubaron con Protein A/G Plus Agarose (Santa Cruz Biotechnology SC-2003) por 30 min a 4°C. Después de 3 lavados de 10 min en PBS, las proteínas precipitadas se analizaron por WB. Los anticuerpos primarios utilizados para la co-IP fueron: anticuerpo policlonal de conejo contra SLX4 (regalo del Dr. J. Rouse, Dundee), anti-XPF (Thermo MS-1381), anti-ERCC1 (Santa Cruz FL297 and Acris-AP17002PU-N) y anti-MUS81 (Abcam ab14387). Los anticuerpos para el análisis de las co-IP por WB fueron anti-XPF (ab76948, Abcam) 1:1000; anti-XPF Ab-1 (Clone: 219, Thermo Scientific-Lab Vision), 1:500; anti-p84 (Abcam ab487), 1:300; anti-Vinculina (ab18058, Abcam), 1:3000. La separación de las fracciones celulares citoplasmática, nuclear y cromatínica se obtuvo mediante la utilización del kit Subcellular Protein Fractionation Kit (Thermo Scientific) siguiendo las instrucciones del fabricante y 25 µg de proteínas de cada fracción se analizaron mediante WB. Para el análisis por WB de las fracciones celulares se utilizaron los siguientes anticuerpos de Abcam: anti-XPF, (ab17798, Abcam) 1:200; anti-tubulina, (ab44928, Abcam) 1:2500; anti-p300, (ab3164, Abcam) 1:200; anti-histona H3, (ab1791, Abcam) 1:3000.

IV.17. Evaluación del desenganche (*unhooking*) por ensayo cometa

La capacidad de *unhooking* de las varias formas mutadas de XPF se midió aproximadamente como se describe en (De Silva et al., 2000, De Silva et al., 2002) mediante electroforesis de células individuales o ensayo cometa (*Comet Assay*) (Azqueta and Collins, 2013). Para inducir los ICL, las células se trataron por duplicado con Melfalán (Sigma) a la concentración indicada en medio de cultivo durante 1 h. A continuación se lavaron dos veces las células con PBS y se dejaron en medio de cultivo fresco a 37°C y 5% de CO₂ por los tiempos indicados. A continuación, se recogieron las células por tripsinización, se resuspendieron a la concentración de 7x10⁵/ml en medio de congelación (90% de FCS, 10% de DMSO) y se guardaron a -80°C hasta el momento del ensayo. Inmediatamente antes de la

electroforesis, las células descongeladas se irradiaron con IR (30 Gy) y 25 μ l de la suspensión celular se añadieron a 225 μ l de una solución de agarosa de bajo punto de fusión (0,75% en PBS sin Ca^{++} y Mg^{++} , EDTA 10 mM). Para cada punto se pipetearon tres gotas de 7 μ l (correspondiente aproximadamente a 500 células) de la suspensión sobre de una membrana GelbondTM (Lonza) donde se les dejó solidificar. Las membranas se pusieron en solución de lisis (1% de Triton-X, 20% de DMSO, 1% de lauril sarcosinato de sodio, NaCl 2,5 M, EDTA 0,1 M, Tris 10 mM, 0,2 M de NaOH, pH 10) por 1 h a 4°C. A continuación se dejaron en tampón de electroforesis fresco (0,3 M de NaOH, EDTA 1 mM, pH 13,2) durante 35 min a 4°C y se realizó la electroforesis durante 20 min a 20 V y 300 mA. Terminada la electroforesis las membranas se dejaron toda la noche en etanol absoluto a 4 °C (fijación). A partir de entonces, se dejaron secar durante 2 h a temperatura ambiente y se rehidrataron durante 20 min en tampón TE pH 7.5 en presencia del fluorocromo SYBRGold (1:10.000, Invitrogen) para teñir el DNA. Las membranas se montaron para el análisis por microscopia de fluorescencia y para cada punto se evaluó el “Olive Tail Moment” (TM) de 200 células con el software Komet 5.5 (Tecnología Andor). El TM se define como el producto de la longitud de la cola y la fracción de DNA total en la cola y, por lo tanto, incorpora una medida del tamaño del fragmento más pequeño detectable de DNA roto (que se refleja en la longitud de la cola del cometa) y el número de piezas rotas (representado por la intensidad de la fluorescencia del DNA en la cola). El porcentaje de unhooking se midió con la fórmula: $(\text{TM}_{mi} - \text{TM}_c)/(\text{TM}_{ci} - \text{TM}_c) \times 100$, donde TM_{mi} es el momento medio de la cola de las muestras tratadas con Melfalán (m) e irradiadas (i), TM_c es la media del momento cola de células sin ningún tratamiento (control, c) y TM_{ci} es el momento medio de la muestra de control (c) sin Melfalán e irradiada (i).

IV.18. Purificación de XPF-ERCC1 y ensayos de reparación *in vitro*

La producción de baculovirus se realizó tal como fue descrito en (Enzlin and Scharer, 2002) utilizando constructos pFastBac1-XPF previamente generados. El XPF-wt, XPF-R689A y p.Arg689Ser se co-expresaron con ERCC1 wt en células de insecto Sf9. Los

heterodímeros se purificaron por afinidad al níquel, cromatografía de exclusión por tamaño y heparina como descrito (Enzlin and Scharer, 2002). Los ensayos de NER con extractos de células XPF deficientes y el plásmido que contiene un aducto de cis-platino en posición conocida en una de las dos cadenas (1,3-intrastrand d (GpTpG)-cisplatin cross-link), se realizaron como anteriormente reportado (Wood et al., 1988, Moggs et al., 1996). Para cada punto analizado, se preparó una reacción en un volumen final de 9 μ l. Cada reacción era constituida por: 2 μ l de tampón de reparación (200 mM Hepes-KOH, 25 mM MgCl₂, DTT 2,5 mM, ATP 10 mM, fosfo-creatina 110 mM, 1,8 mg/ml de BSA, pH final 7,8); 0,2 μ l de tampón de Creatin-fosfoquinasa (2,5 mg/ml Creatin-fosfoquinasa de músculo de conejo, (Sigma), Glicina 10 mM, pH 9,0, 50% de Glicerol); 3 μ l de extracto de células deficientes en XPF; NaCl 70 mM final; proteínas XPF purificadas (33,5 nM). Cada reacción se calentó a 30°C durante 10 min. Se añadió 1 μ l (50 ng/ml) de plásmido con la lesión a cada reacción y se incubaron a 30°C durante 45 min. La lesión presente en el plásmido se repara específicamente por NER y el éxito de la reparación se puede verificar por la presencia de fragmentos de DNA de 24-32 nucleótidos de longitud que son escindidos a lo largo del proceso de reparación. Estos fragmentos se detectan por hibridación con un oligonucleótido complementario que contiene una secuencia de 4 Guanosinas que sobresalen en 5'. Una vez que se haya formado el oligonucleótido de dsDNA se rellena la segunda cadena con α -³²P-dCTP. Para este propósito, la mezcla de reacción se enfrió en hielo durante 5 min, seguido por adición de 0.5 μ l de una solución 1 mM de la siguiente cadena complementaria:

5' GGGGGAAGAGTGCACAGAAGAAGACCTGGTTCGACCp 3'.

La mezcla de reacción se desnaturalizó a 95°C durante 5 min y después de 15 min a temperatura ambiente, se añadió 1 μ l de mezcla con 0.13 unidades de Sequenasa y 2 μ Ci de [α -³²P] dCTP a cada reacción. Después de una pre-incubación a 37°C durante 3 min, se añadió 1.2 μ l de mezcla de dNTP (50 μ M dCTP, 100 μ M dTTP, 100 μ M dATP and 100 μ M dGTP). Se incubaron las muestras a 37°C durante 12 min y se detuvo la reacción mediante la adición de 8 μ l de tampón carga (formamida al 80% y EDTA 10 mM). Después de 5 min a 95°C, las muestras se cargaron en un gel desnaturalizante al 14% de poliacrilamida. Los geles se corrieron a 45 W durante 1 h y las bandas se visualizaron en un aparato PhosphorImager Typhoon 9400, (Amersham Biosciences).

Para el ensayo de actividad nucleasa se calentaron a 95°C por 10 min en 200 µl de una solución de annealing (10 mM Tris pH 8.0, 50 mM NaCl, 1 mM MgCl₂) 10 pmol del oligonucleótido con secuencia GCCAGCGCTCGG(T)₂₂CCGAGCGCTGGC marcado con el fluorocromo Cy5 en la extremidad 3'. La solución se dejó enfriar lentamente hasta la temperatura ambiente a lo largo de dos horas. Para cada reacción se utilizaron 100 fmol de producto de annealing. La solución de reacción para el ensayo estaba compuesta por 25 mM Tris pH 8.0, 2 mM MgCl₂ o 0.4 mM MgCl₂, 10% glicerol, 0.5 mM β-mercaptoetanol, 0.1 mg/ml BSA, 40 mM NaCl y 100 fmol de las diferentes proteínas en un total volumen de 15 µl. Se incubó a 30°C por 30 min y se detuvo la reacción mediante la adición de 8 µl de tampón de carga (formamida al 80% y EDTA 10 mM). Después de 5 min a 95°C y enfriamiento en hielo, las muestras se cargaron en geles desnaturalizantes al 12% de poliacrilamida. Los geles se corrieron a 50°C por 40 min y las bandas se visualizaron en un aparato PhosphorImager Typhoon 9400 (Amersham Biosciences).

RESULTADOS

“That's funny.”

Alexander Fleming

V. RESULTADOS

V.1. Subtipaje y caracterización molecular de los pacientes FA mediante método WES + MLPA

49 pacientes de FA diagnosticados en base a la sintomatología clínica y por el ensayo de fragilidad cromosómica se sometieron a análisis por WES. De estos pacientes 45 son de origen española y 4 (ING1, ING2, ING3 y ING4) son de origen del Reino Unido. 34 no presentaban ningún fenómeno de mosaicismo (Full-FA) mientras que 15 eran mosaicos. En 30 pacientes se había ya determinado por subtipaje retroviral el grupo de complementación mientras que en 10 pacientes solo se habían descartado algunos subtipos pero no se había logrado encontrar el grupo de complementación y 9 pacientes estaban totalmente para subtipar. Todas las mutaciones encontradas por WES se confirmaron sucesivamente por secuenciación del DNA genómico por el método Sanger.

V.1.1. Análisis de la cobertura de los genes FA en el análisis por WES.

Para la creación y el enriquecimiento de las librerías para el análisis por WES de los pacientes FA se utilizó el kit comercial SureSelectXT All Exon, V5, 50.621.018 bp de Agilent específico para todo el exoma. Para los pacientes FA104 y FA287 analizados por sistema Solid se utilizó el kit de captura “SureSelect Target Enrichment Human All Exon” de Agilent (véase materiales y métodos). Hasta al momento no se conocía el grado de cobertura que tenían los genes FA en el análisis por WES cuando se utilizaban kits de captura comerciales para el exoma entero. Uno de los primeros pasos fue, por tanto, confirmar que todos los exones de los genes FA tuvieran suficiente cobertura para asegurarnos que utilizar

la WES como aproximación para el diagnóstico molecular de los pacientes fuera viable. En la Figura 13 se puede observar las coberturas de los 18 genes FA identificados hasta el momento en 41 de los pacientes de nuestra cohorte y la cobertura de todos los genes FA en cada uno de 41 de los 49 pacientes analizados (Figura 13B).

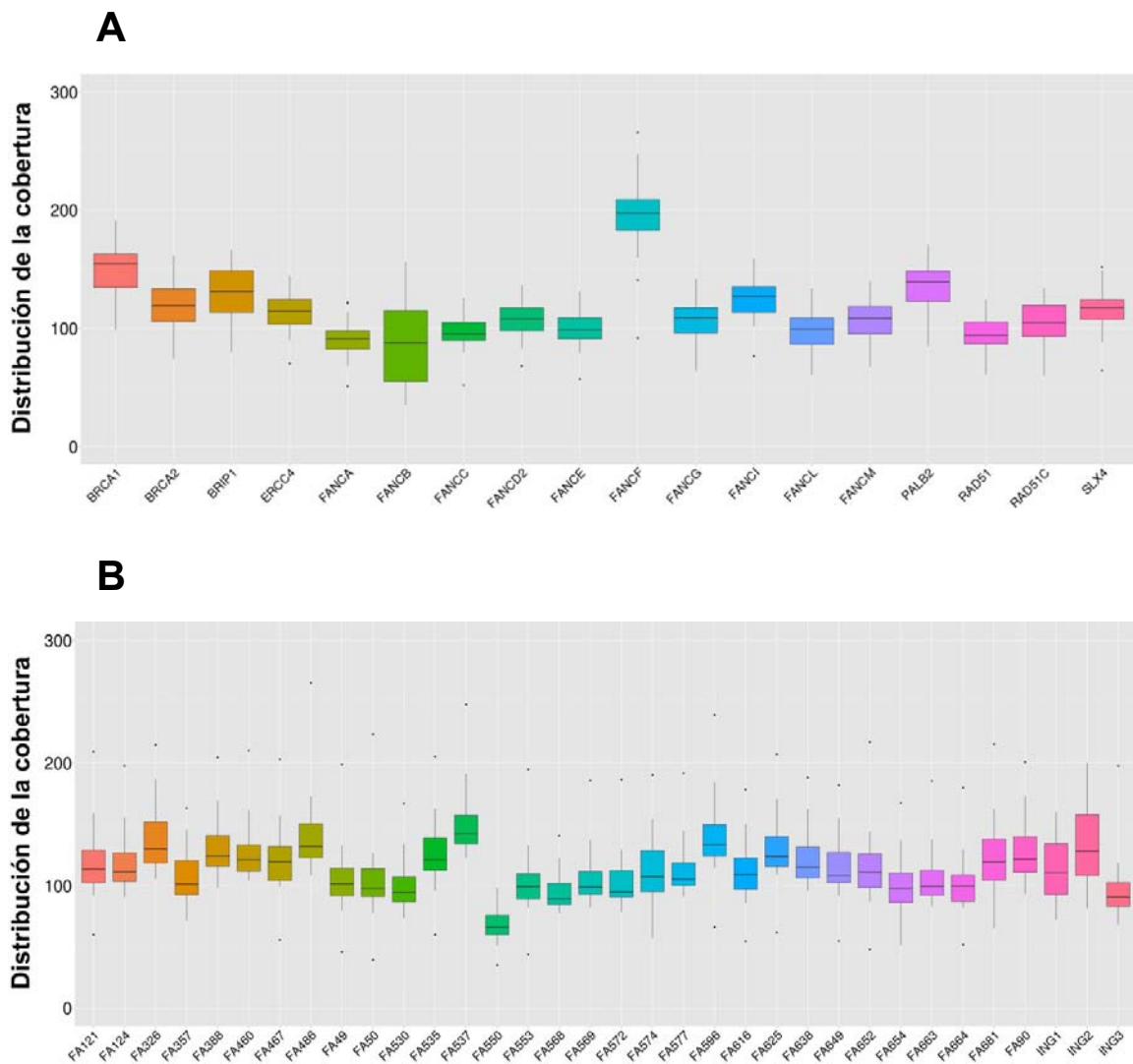


Figura 13: (A): Cobertura media por gen FA en la cohorte de pacientes y (B): Cobertura media por cada paciente de todos los genes FA.

Todos los genes FA tienen unas coberturas medias de alrededor de 100x y *FANCF* alrededor de 200 y el valor de cobertura más bajo encontrado fue de 35x para el gen *FANCB*. Hay que recordar que el gen *FANCB* se encuentra en el cromosoma X y que por tanto solo

hay un alelo de *FANCB* en los varones de nuestra cohorte. Considerando que para una identificación fiable de las variantes en los estudios de las enfermedades genéticas raras por WES se recomienda una cobertura por lo menos de 30 (Meynert et al., 2013), estos datos indican claramente que es posible utilizar con confianza la aproximación por WES para la caracterización molecular de los pacientes FA utilizando un kit de captura comercial.

V.1.2. Análisis por WES y MLPA de 49 pacientes FA

Para la identificación de las mutaciones se utilizó como criterio que fuera una mutación ya reportada como patogénica (*Fanconi Anemia Mutation Database*, <http://www.rockefeller.edu/fanconi/>), que fuese una mutación claramente patogénica (*frameshift*, *nonsense* y *splicing*) aunque no estuviese reportada antes y que la variante tuviese una prevalencia alélicas (*Minor allele Frequency*; MAF) < 0.01. Como ya se ha explicado en la introducción, alrededor de un 15-20% de pacientes FA-A tienen grandes deleciones (Morgan et al., 1999, Ameziane et al., 2008, Gille et al., 2012) en el gen *FANCA*. El análisis por WES solo detecta cambios de nucleótidos o pequeñas inserciones y deleciones de un número limitado de bases, por lo tanto analizamos por MLPA los pacientes FA-A donde no se consiguió encontrar una o ambas mutaciones por WES.

En el caso los pacientes Full-FA (Tabla VI) acoplando WES y MLPA fuimos capaces de identificar las dos mutaciones en 31 de los 34 pacientes Full-FA (91,2%); en 3 pacientes, (2 subtipados FA-A y 1 no subtipado) identificamos solo una mutación en el gen *FANCA*. Cinco pacientes (FA287, FA326, FA357, FA531, FA569) presentaban mutaciones en genes (*FANCI*, *FANCG*, *FANCD2*, *FANCD2* y *FANCA* respectivamente) que el subtipaje tradicional había descartado. En el paciente FA572 identificamos una sola mutación en el gen *FANCA* que había sido descartado por los estudios de complementación retroviral. El paciente FA104 resultó tener mutaciones bialélicas en el gen *ERCC4* que nunca se había encontrado mutado en un paciente de FA. En este grupo de pacientes no mosaico identificamos 65 alelos mutados sobre de un total de 68 alelos (95,6%) y todas las variantes se confirmaron por secuenciación del DNA de Sanger. En los pacientes mosaicos (Tabla VI) logramos identificar todas las mutaciones causantes de la enfermedad en 11 de los 15 pacientes (73.3%) mientras que en 4 pacientes (FA121, FA664, FA681 y FA124) pudimos

identificar solo una mutación en los genes *FANCE* (FA121) y *FANCD2* (FA664, FA681 y FA124). En este grupo de pacientes conseguimos identificar 25 de los 29 alelos de genes FA mutados en nuestra población (86.2%). Los alelos son 29 y no 30 porque el paciente FA574 es varón y tiene una mutación en el único alelo de *FANCB*. Hay que remarcar que para la caracterización por WES de los 4 pacientes donde se identificó solo una de las mutaciones se utilizó DNA de PB y que la MLPA no se ha implementado para el análisis de deleciones en los genes *FANCE* y *FANCD2*.

Resumiendo, utilizando la aproximación WES + MLPA fuimos capaces de caracterizar molecularmente por completo a 42 de los 49 pacientes analizados (85,7%). De estos, 31 pacientes eran FA-A, 1 FA-B, 2 FA-D1, 3 FA-D2, 3 FA-G 1 FA-J y 1 FA-Q. Si se consideran los alelos, en nuestra población de pacientes identificamos 90 de los 97 alelos mutados (92,8%). En 10 pacientes (FA569, FA357, FA287, FA104, FA331, FA531, FA326, FA574, FA572, FA124) se habían descartado varios genes utilizando el método de subtipaje retroviral. En 6 de estos pacientes encontramos ambas mutaciones en genes que se habían descartados (FA569, FA357, FA287, FA331, FA531, FA326). De los cuatro pacientes restantes, 2 (FA124 y FA574) presentaban una sola mutación en genes diferentes a los que se habían descartado y 1 paciente (FA572) presentaba una mutación en un gen que había dado negativo en el subtipaje tradicional. En el último paciente (FA104) se encontraron dos mutaciones en el gen *ERCC4* que nunca se había encontrado hasta al momento mutado en la FA. Por lo tanto 7 de los 10 pacientes en que los estudios por complementación retroviral habían descartado subtipos sin identificar el de pertenencia, resultaron ser “falsos negativos”. Hay que subrayar que no se encontró ningún “falso positivo” por subtipaje retroviral. De los 10 pacientes mosaicos que para su caracterización se utilizó DNA proveniente de PB, en 4 se encontró una sola mutación (40%) mientras que se identificó solamente una mutación en 3 muestras de DNA de PB de 29 pacientes Full-FA (10,3%). Esto parece indicar que para la caracterización molecular de los pacientes mosaicos, el DNA proveniente de PB probablemente no es el más adecuado para el análisis por WES + MLPA.

En nuestra cohorte de pacientes se identificaron 22 mutaciones que no se habían reportado anteriormente en la literatura y que no estaban catalogadas como SNPs. De estas 22 mutaciones, 17 eran mutaciones claramente patogénicas (nonsense, frameshift y splicing) mientras que 5 eran de cambio de sentido (Tabla VII). Las mutaciones de cambio de sentido

se analizaron con las herramientas Polyphen-2 (Adzhubei et al., 2010) y SIFT (Kumar et al., 2009) y con ambas aproximaciones resultaron tener una alta probabilidad de ser patogénicas.

V.1.3. Análisis por cobertura de las grandes deleciones en FANCA

Durante la WES, cada una de las bases que forman los segmentos de DNA de una librería se secuencian varias veces en reacciones independientes. El número de veces que una base se secuencian constituye el dato de la cobertura (*coverage*) para dicha base. Si por efecto de una deleción intragénica uno o más exones de un gen no están presentes, la cobertura de las bases que pertenecen a los exones afectados tendrá una variación significativa respecto a los controles: si la deleción está presente en heterocigosis la cobertura será reducida de un 50% mientras que si está en homocigosis la cobertura de todas las bases incluidas en esta región será igual a cero. En el caso de las deleciones en heterocigosis, cuanto más alto es el valor de la cobertura, más fiables serán las diferencias a la hora de determinar la variación del número de copias (*copy number variation*; CNV). Este tipo de aproximación ya se ha empleado para determinar CNV a partir de datos de NGS generados con paneles de genes (donde la cobertura puede llegar hasta 1000x (Jamuar et al., 2014, Ameziane et al., 2012)) pero nunca se ha utilizado a partir de datos obtenidos con un kit de captura comercial. Consecuentemente, quisimos verificar si utilizando el dato de la cobertura se podían detectar las deleciones intragénicas en nuestros pacientes estudiados por WES. En la Figura 14 se puede observar el análisis con el software IGV (*Integrative Genomics Viewer*) (Robinson et al., 2011) de los datos de WES de dos pacientes de nuestra cohorte en comparación con el análisis por MLPA. El paciente ING2 (Figura 14 A, en rojo) no presenta señal de cobertura para los exones que van desde el 16 hasta el 20 de *FANCA* confirmando así el dato de la MLPA. El paciente FA553 (Figura 14 B, en rojo) presenta una cobertura reducida a la mitad para los exones 32 y 33 con respecto a muestras de control. Esto es en línea con el resultado de la MLPA que muestra una deleción en heterocigosis de los exones 32 y 33 de *FANCA*. En la MLPA del paciente FA553 se observa también una deleción en heterocigosis del exón 13. Esto es debido a que el otro alelo de *FANCA* presenta una microdeleción (c.1115_1118delTTGG) que afecta a la secuencia de hibridación de uno de los cebadores específicos para el exón 13 de *FANCA*.

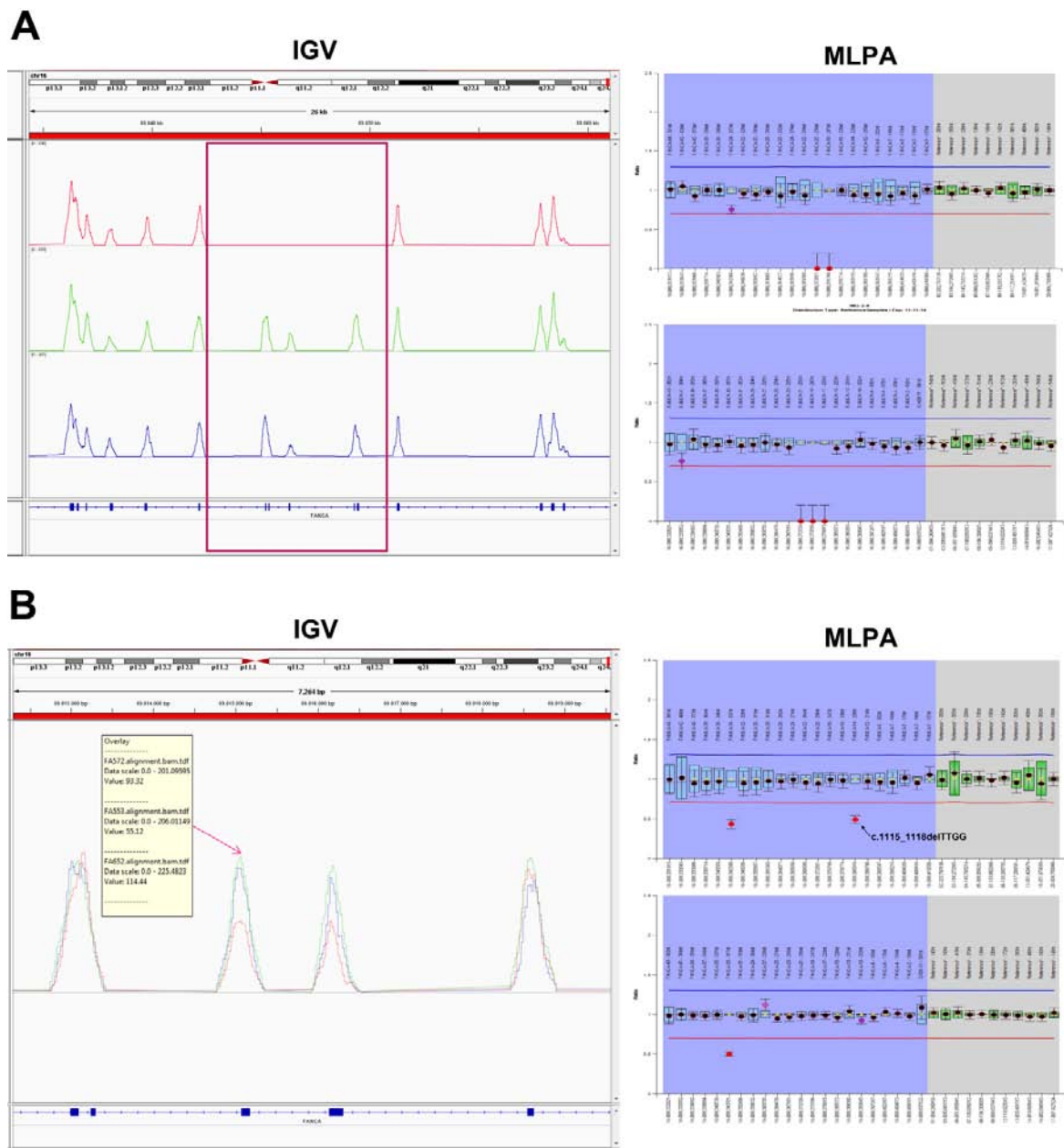


Figura 14: (A) Análisis por cobertura de una deleción intragénica en homocigosis: el paciente ING2 (IGV, en rojo) no presenta señal de cobertura para los exones desde el 16 hasta el 20 de FANCA en comparación con muestras analizadas en la misma tanda. (B) Reducción del 50% de la cobertura en la muestra FA553 (IGV, en rojo) para los exones 32 y 33 indicativa de una deleción intragénica en heterocigosis. En el panel MLPA se evidencia la microdeleción c.1115_1118delTTGG que afectando al sitio de hibridación de uno de los cebadores aparece como una deleción del Exón 13 de FANCA.

Tabla VI: Mutaciones identificadas en los pacientes FA

FULL-FA		3. Mutación		3. Mutación	
CODIGO	Sexo	Subtipo	DNA	Gen	Mutación
FA104	F	No A,B,C,G,E,F,D2,J, Foci Rad51	PB	FANCD2/BRCA2	Exón 11; c.2065C>A; p.Arg688Ser; Bogliolo, 2013
FA156	F	A	LCL	FANCA	Exón 36; c.3558dupG; p.Arg1186Glnfs*28; Savino, 2003
FA287	F	No A,B,C,G,E,F,D2,J, Foci Rad51	PB	FANCI	Exón 7; c.751C>T; p.Arg251Cys; Chandrasekharappa,2013
FA326	F	No A,C,G,E,F,D2	PB	FANCG	Exón 13; c.1642C>T; p.Arg548*; Ahuerbach, 2003
FA357	F	No A,B,C,G,E,F,D2	PB	FANCD2	Exón 3; c.98dupA; p.Lys34Glnfs*6; No reportada
FA383	F	D2	PB	FANCD2	Exón 26; c.2444G>A; p.Arg815Gln; Kalb, 2007
FA388	F	A	PB	FANCA	Exón 13; c.1115_1118delTTGG; p.Val372Alafs*42; FA/Breast cancer consortium, 1996
FA460	F	A	PB	FANCA	Exón 37; c.3763G>T; p.Gln1255*; No reportada
FA486	F	G	PB	FANCG	Exón 13; c.1642C>T; p.Arg548*; Ahuerbach, 2003
FA522	M	G	PB	FANCG	Exón 4; c.450G>A; p.Trp150*; No reportada
FA530	F	ND	PB	FANCA	Exón 36; c.3558insG; p.Arg1186Glnfs*28; Savino, 2003
FA531	F	No A,C,G,E,F,D2	PB	FANCD2	Exón 26; c.2444G>A; p.Arg815Gln; Kalb, 2007
FA535	M	A	PB	FANCA	Exón 13; c.1115_1118delTTGG; p.Val372Alafs*42; FA/Breast cancer consortium, 1996
FA537	F	A	PB	FANCA	Exón 38; c.3788_3790delTCT; p.Phe1263del; Magdalena, 2005
FA551	M	A; Hermano/a del FA550	PB	FANCA	Exón 38; c.3788_3790delTCT; p.Phe1263del; Magdalena, 2005
FA559	F	No A, B, C, puede G	PB	FANCA	Exón 8; c.790C>T; p.Gln264*; Savino, 1997
FA572	F	No A,C,G,D2,E,F	PB	FANCA	Exón 27; c.2574C>G; p.Ser858Arg; Tamary, 2000
FA577	F	A	PB	FANCA	Delección del Exón 3
FA596	M	A	PB	FANCA	Exón 13; c.1115_1118delTTGG; p.Val372Alafs*42; FA/Breast cancer consortium, 1996
FA616	M	A	PB	FANCA	Exón 2; c.165C>T; p.Gln55*; Wilker, 1999
FA638	F	A	PB	FANCA	Exón 12; c.1039G>T; p.Gln345*; No reportada
FA649	M	A	PB	FANCA	Exón 28; c.2641C>T; p.Gln881*; Castilla, 2011b
FA652	M	A	PB	FANCA	Exón 37; c.3763G>T; p.Gln1255*; No reportada
FA654	M	A	PB	FANCA	Exón 21; c.1860dupC; p.Tyr621Leufs*9; No reportada
FA705	M	ND	PB	FANCA	Exón 39; c.3913C>T; p.Leu1305Phe; Rockfeller DB
FA707	M	ND	PB	FANCA	Exón 13; c.1115_1118delTTGG; p.Val372Alafs*42; FA/Breast cancer consortium, 1996
FA712	M	ND	PB	FANCA	Exón 35; c.3283C>T; p.Ser1088Phe; Rockfeller DB
FA718	F	ND	PB	FANCA	Exón 13; c.1115_1118delTTGG; p.Val372Alafs*42; FA/Breast cancer consortium, 1996
FA746	F	ND	PB	FANCA	Exón 23; c.2128A>T; p.Arg710*
ING11	ND	ND	PB	FANCA	Exón 38; c.3788_3790delTCT; p.Phe1263del; Magdalena, 2005
ING21	ND	A	PB	FANCA	Exón 29; c.2837_2838insT; p.Ser947Phefs*4*; No reportada
ING31	ND	A	PB	FANCA	Delección Exones 16-20
ING41	ND	A	PB	FANCA	Delección Exones 1-6
					Delección Exones 18-21

†: Pacientes originarios de Reino Unido

MOSAICOS		3. Mutación		3. Mutación	
CODIGO	Sexo	Subtipo	DNA	Gen	Mutación
FA049	M	A; Hermano/a FA050	LCL	FANCA	c.1083>21>C; splicing; No reportada
FA050	M	A; Hermano/a FA049	LCL	FANCA	c.1083>21>C; splicing; No reportada
FA090	F	A	LCL	FANCA	Exón 36; c.3558insG; p.Arg1186Glnfs*28; Savino, 2003
FA121	M	E	PB	FANCE	Exón 4; c.929insC; p.Val311Serfs*1; no reportada
FA124	F	No A/C/G	PB	FANCD2	Exón 13; c.1068T>A; p.Tyr356*; no reportada
FA158	M	ND	PB	FANCD2/BRCA2	c.8488-1G>A; IVS19:1 G to A; Howlett, 2002
FA331	M	No A,C,G	PB	FANCA	Exón 40; c.5982A>G; p.Trp1528Ala; Rockfeller DB
FA550	F	A; Hermano/a del FA551	PB	FANCA	Exón 38; c.3788_3790delTCT; p.Phe1263del; Rockfeller DB
FA553	M	A	PB	FANCA	Exón 13; c.1115_1118delTTGG; p.Val372Alafs*42; FA/Breast cancer consortium, 1996
FA568	F	A	PB	FANCB	Exón 39; c.3913C>T; p.Leu1305Phe; Rockfeller DB
FA574	M	No A,C,G; no ubiquitina FANCD2	PB	FANCB	Exón 3; c.786G>T; p.Lys262Asn
FA609	M	A	PB	FANCA	Exón 1 c.65G>A; p.Trp22*
FA653	F	ND	PB	FANCD2/BRCA2	Exón 16; c.7756A>G; p.Glu2595Gln; no reportada
FA664	F	ND; Hermano/a FA681	PB	FANCD2	Exón 3; c.226_228delA; p.Lys77Rfs*7; no reportada
FA681	M	ND; WB; No FANCD2; Hermano/a FA664	PB	FANCD2	Exón 3; c.226_228delA; p.Lys77Rfs*7; no reportada

Tabla VII: Mutaciones no previamente reportadas.

Paciente	Gen	Mutación	Polyphen-2	SIFT
FA049	FANCA	c.1083+2T>C; splicing	NA	NA
FA090	FANCA	c.472insC; p.Hist158Profs*23	NA	NA
FA104	FANCC/ERCC4	c.2065C>A; p.Arg689Ser	Probably damaging	Damaging
FA104	FANCC/ERCC4	c.1484_1488delCTCAA; p.Thr495Asnfs*5	NA	NA
FA121	FANCE	c.929insC; p.Val311Serfs*1	NA	NA
FA124	FANCD2	c.1068T>A ; p.Tyr356*	NA	NA
FA287	FANCC	c.975_977delA;p.Arg279Glyfs*10	NA	NA
FA326	FANCC	c.566dupA; p.Glu190Glyfs*45	NA	NA
FA357	FANCD2	c.98dupA; p.Lys34Glyfs*6	NA	NA
FA460	FANCA	c.3763G>T; p.Glu1255*	NA	NA
FA522	FANCC	c.450G>A; p.Trp150*	NA	NA
FA535	FANCA	c.1100G>T; p.Ser367Ile	Possibly Damaging	Damaging
FA574	FANCB	c.786G>T; p.Lys262Asn	Probably damaging	Damaging
FA625	FANCA	c.1033G>T; p.Glu345*	NA	NA
FA638	FANCA	c.3448dupC; p.Leu1150Profs*65	NA	NA
FA649	FANCA	c.3762G>T; p.Glu1255*	NA	NA
FA649	FANCA	c.3335T>G; p.Val1112Gly	Probably damaging	Damaging
FA652	FANCA	c.1860dupC ; p.Tyr621Leufs*9	NA	NA
FA652	FANCA	c.893+2T>C ; Splicing	NA	NA
FA663	FANCD1/BRCA2	c.7796A>G; p.Glu2599Gly	Probably damaging	Damaging
FA664	FANCD2	c.226_228delA; p.Lys777Argfs*7	NA	NA
ING1	FANCA	c.2837_2838insT; p.Ser947Phefs4*	NA	NA

V.2. Mutaciones bialélicas en *ERCC4* causan FA

El paciente FA104 era un típico paciente FA, con BMF, los pulgares ausentes y otras malformaciones típicas de la enfermedad. La prueba de fragilidad cromosómica inducida por DEB confirmó sin ambigüedades el diagnóstico de FA con un 92% de células que presentaban roturas cromatídicas con un promedio de 4,44 roturas por célula. Los datos de la fragilidad cromosómica de este paciente se muestran en la Figura 15 A en el marco de una serie histórica de datos de pacientes españoles donde se examinó la fragilidad cromosómica para confirmar el diagnóstico de FA. Los linfoblastos FA104 eran hipersensibles a la MMC, mostraban monoubiquitinación normal de FANCD2 y no presentaban sensibilidad ni a los inhibidores de la Topoisomerasa I (Figura 13B), ni a los inhibidores de la PARP (Figura 13C) y formaban *foci* normales de RAD51 en respuesta al daño en el DNA (Casado et al., 2007).

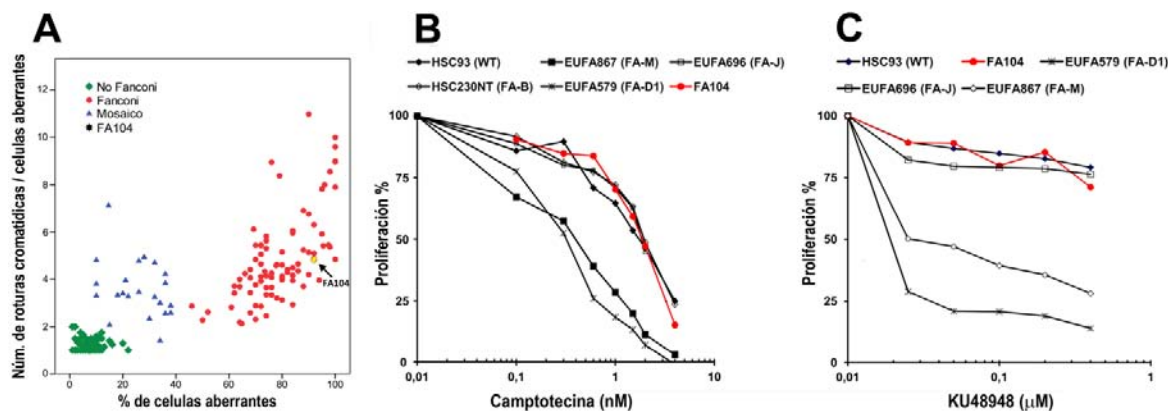


Figura 15: (A) Fragilidad cromosómica después de tratamiento con 0,1 μg/ml de DEB en el paciente FA104 en el marco de la base de datos española de fragilidad cromosómica de la FA (Castella et al., 2011a). (B) La LCL FA104 (en rojo) no muestra una sensibilidad a los inhibidores de la Topoisomerasa I (Camptotecina) o (C) de la PARP (KU58948) en comparación con LCL wt (HSC93) o con mutaciones en *BRIP1/FANCI* (EUFA696), *BRCA2* (EUFA579) o *FANCM* (EUFA867)

Esto sugería que el defecto en la vía FA/BRCA estaba por debajo de la monoubiquitinación de FANCD2 y que no implicaba alteraciones de BRCA2 o de la HRR pero todos los intentos realizados hasta al momento para identificar el gen mutado no habían tenido éxito. En un primer análisis de los datos generados por WES, se verificó que no había habido errores en el subtipaje retroviral y se descartaron todas los SNPs ya reportados en la base de datos de “1000 Genomas” y las variantes con una cobertura menor de 9 por ser poco fiables. Al no presentarse la enfermedad en los padres y siendo el paciente FA104 de sexo femenino, los resultados se cribaron en base a un modelo de herencia autosómica recesiva. Por lo tanto, se consideraron como posibles candidatos los genes que tenían por lo menos dos variantes no sinónimas nunca reportadas anteriormente. Siguiendo este modelo, identificamos 16 genes que presentaban por lo menos dos mutaciones como posibles candidatos a ser causa de la FA en el paciente FA104 (Tabla VIII). De los genes en la lista, *ERCC4* inmediatamente nos llamó la atención por ser ya conocido como un gen involucrado en la reparación del daño en el DNA. La proteína XPF, codificada por *ERCC4*, es el componente catalítico de una endonucleasa involucrada en la reparación del DNA. XPF forma con la proteína ERCC1 un dímero que es la endonucleasa responsable para la incisión en 5' del DNA durante la NER (Tripsianes et al., 2005, Ciccia et al., 2008).

Tabla VIII: Lista de los genes candidatos en el paciente FA104:

Chrom	Pos	Ref	Alt	Ensembl pred.	AA	Gen	NRR	SNV Q	GT Q
1	169489751	A	W	SS	-	F5	42	171	171
1	169525877	T	Y	SS	-	F5	52	36	36
2	73675227	-	CTC	NFC	S/SP	ALMS1	16	N/A	N/A
2	73678183	G	R	NSC	G1509D	ALMS1	156	120	120
3	49094490	G	S	NSC	N381K	QRICH1	122	228	228
3	49095011	C	S	NSC	G208R	QRICH1	109	43	43
4	126238305	C	M	NSC	P247T	FAT4	52	178	178
4	126355484	C	M	NSC	A2368E	FAT4	56	190	190
5	156479444	TTG	-	NFC	TS/S	HAVCR1	61	N/A	N/A
5	156479568	-	GTT	NFC	T/TT	HAVCR1	106	N/A	N/A
6	31238942	G	W	NSC	A176V	HLA-C	23	61	39
6	31239577	A	C	NSC	S48A	HLA-C	21	90	90
6	32709309	A	R	SS	-	HLA-DQA2	29	84	84
6	32713044	C	Y	NSC	T64M	HLA-DQA2	192	228	228
6	32713188	C	Y	SS	-	HLA-DQA2	126	228	228
6	38840915	A	R	NSC	I2479V	DNAH8	72	216	216
6	38879340	A	T	NSC	E3267D	DNAH8	12	34	34
7	100686777	C	Y	NSC	T4027M	MUC17	323	228	228
7	100687107	G	R	SS	-	MUC17	66	79	79
8	30700598	T	Y	NSC	N1979S	TEX15	33	97	97
8	30701995	A	M	NSC	D1513E	TEX15	141	228	228
10	69682773	T	Y	NSC	D920G	HERC4	64	69	69
10	69785435	-	A	SS	-	HERC4	9	N/A	N/A
16	14029271	AACTC	-	FC	-	ERCC4	22	N/A	N/A
16	14041518	C	M	NSC	R689S	ERCC4	121	228	228
16	72137553	C	S	NSC	Q564E	DHX38	56	85	85
16	72142141	A	R	NSC	S994G	DHX38	52	106	106
17	74272839	C	Y	NSC	V1593M	QRICH2	54	33	33
17	74277009	T	Y	NSC	Q1264R	QRICH2	23	81	81
18	14105016	C	M	NSC	R508I	ZNF519	136	228	228
18	14105853	C	M	NSC	R229I	ZNF519	23	51	51
19	51918360	A	R	NSC	S445P	SIGLEC12	43	39	39
19	52004795	G	CT	FC	-	SIGLEC12	19	N/A	N/A

Chrom: número del cromosoma; Pos: coordinada cromosómica de la variante (GRCh37/hg19); Ref: alelo de referencia; Alt: alelo encontrado en la muestra; Ensembl pred: Predicción de la consecuencias sobre del transcrito (Ensembl v59). SS=sitio de splicing, NSC=non-synonymous coding, FC=frameshift coding, NFC=non-frameshift coding; AA change: cambio de aminoácido; Gen: Nombre del gen; NRR: número de lecturas no redundantes (cobertura); SNV Q: La probabilidad (phred-scale) que el genotipo sea igual a la referencia; GT Q: Probabilidad (Phred-scale) que el genotipo sea erróneo.

Mutaciones en el gen *ERCC4* son causa de XP en los pacientes pertenecientes al grupo de complementación F (Sijbers et al., 1996, Bessho et al., 1997, Gregg et al., 2011). Alteraciones de XPF (o de ERCC1) como las de las otras proteínas involucradas en la NER, producen sensibilidad a los UVC, pero únicamente las alteraciones de XPF (y de ERCC1)

son causa también de una peculiar y marcada sensibilidad a los ICL (Thompson et al., 1981, Gregg et al., 2011). Ambas variantes en *ERCC4* del paciente FA104 se confirmaron por Sanger (Figura 16A) y se predecían ser patogénicas. Una de las variantes era una delección de 5 bp en el exón 8 (c.1484_1488delCTCAA) que producía un desplazamiento del marco de lectura (p.Thr495Asnfs*5) y daba origen a una proteína XPF truncada sin la mitad carboxi terminal y, por tanto, del dominio catalítico (Ciccia et al., 2008). La otra era una variante de cambio de sentido en el exón 11 (c.2065C>A; p.Arg689Ser) que alteraba un residuo de Arginina altamente conservado en el sitio catalítico de XPF (Enzlin and Scharer, 2002) (Figura 14B). La secuenciación de Sanger del DNA de los padres demostró el origen materno de la delección y el origen paterno de la variante de cambio de sentido. Ninguna de estas variantes se encontró en una población de 200 individuos sanos españoles.

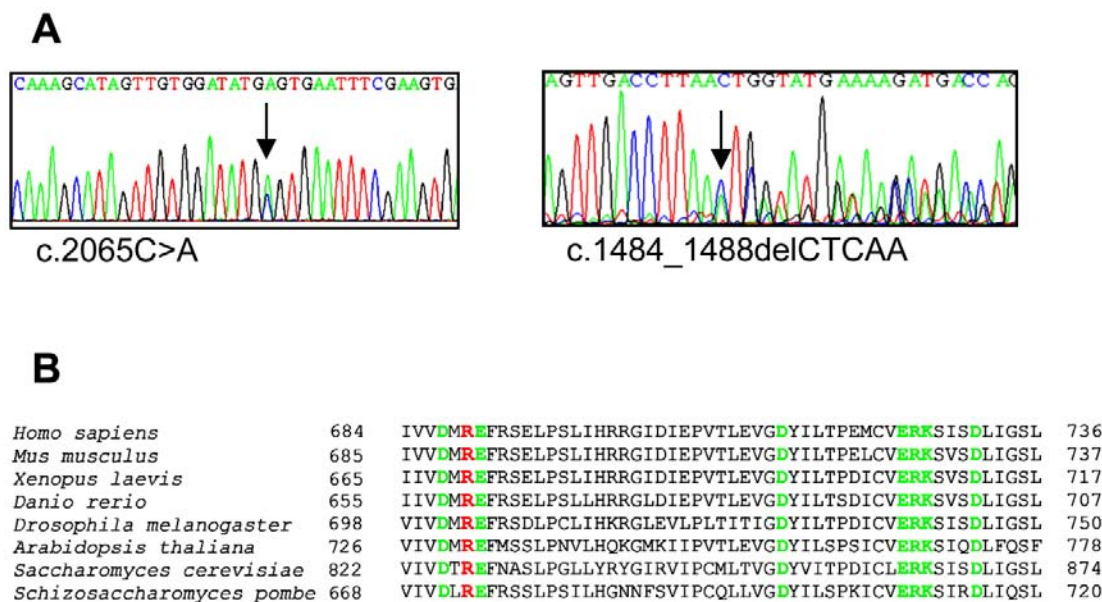


Figura 16: (A) Mutaciones en el gen *ERCC4* en el paciente FA104. **(B)** Conservación del residuo Arg689 en la proteína XPF. Los Aminoácidos que juntos con la Arg689 son fundamentales para la actividad del sitio catalítico de la nucleasa XPF se muestran en verde.

A través de la exposición constante a una dosis sub-tóxica de MMC (véanse Materiales y Métodos) se generó una línea derivada de la FA104 que había revertido su sensibilidad a la MMC (Figura. 17A).

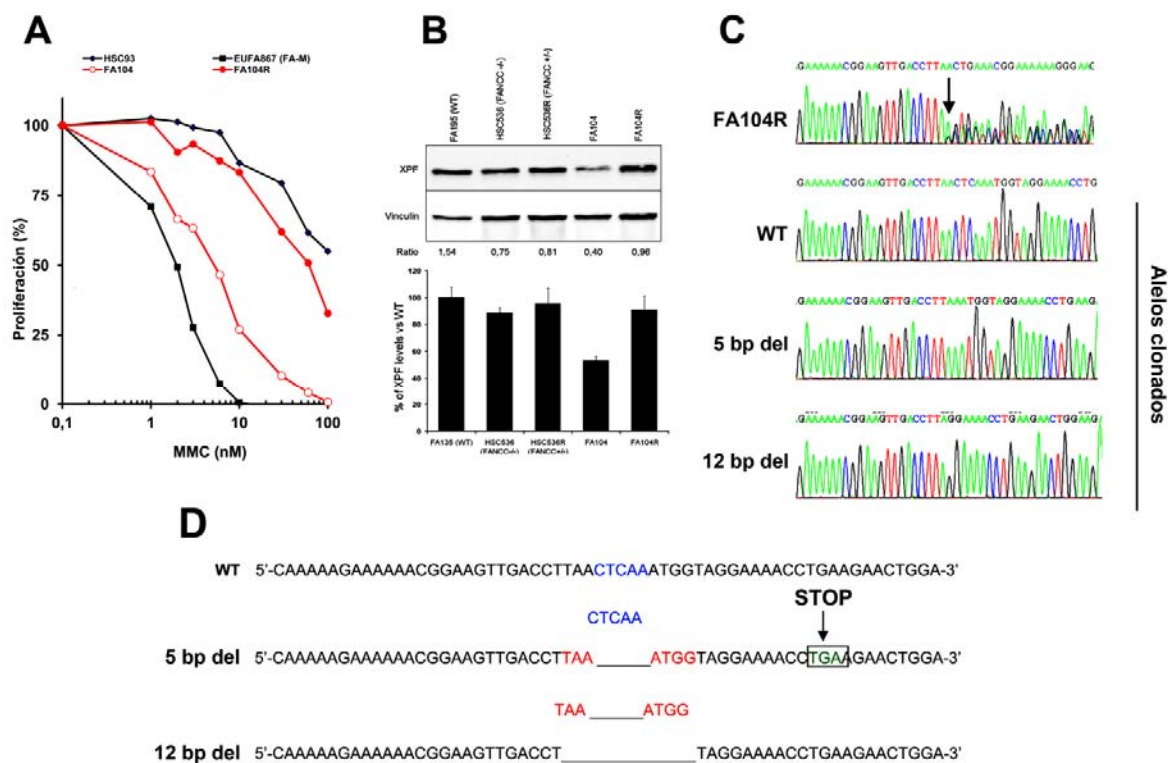


Figura 17: (A) Sensibilidad de la línea revertida FA104R a la MMC. (B) Cuantificación de los niveles de XPF por WB en las LCLs FA104, FA104R, HSC536 (FA-C), HSC536R (HSC536 revertida) and FA139 (wt). Los niveles de XPF se muestran normalizados a los niveles del control de carga (vinculina). En la gráfica se presentan la media y el SEM de tres experimentos independientes. (C) Análisis de los diferentes alelos clonados de la línea FA104R. La secuenciación de las colonias bacterianas individuales revela la presencia de una delección de 12 bp en el exón 8 que se solapaba a la delección de 5 bp y que restauraba la pauta de lectura del gen *ERCC4*. (D) Esquema de como la delección compensatoria de 7 bp (en rojo) restablece el marco de lectura correcto del gen *ERCC4* en la línea FA104R.

El análisis por WB reveló que en la línea revertida (FA104R) los niveles de la proteína XPF habían vuelto a lo normal (Figura 17 B) pero el análisis por secuenciación mostraba que tanto la mutación de cambio de sentido como la mutación de frameshift en *ERCC4* seguían estando presentes en la línea FA104R. Mirando pero más atentamente a los patrones de los picos de la mutación de frameshift, notamos como el punto en el exón 8 de *ERCC4* desde que empezaba el solapamiento de los picos correspondientes a las secuencias de los diferente alelos empezaba tres bases antes que en la FA104 y que se podía notar para algunos picos la presencia de las señales de tres distintos nucleótidos (Figura 17 C). Por lo tanto decidimos amplificar y clonar el producto de PCR mediante TOPO cloning en bacterias (véase Materiales y Métodos) con el fin de poder separar los diferentes alelos de *ERCC4*

presentes en la línea FA104R. Se secuenció el DNA plasmídico de varias colonias independientes y se pudo averiguar que en una subpoblación de la línea FA104R se había producido una segunda deleción de 7 bp en *cis* con la deleción de 5 bp (Figura 17 C). El resultado de esta segunda deleción, que abarcaba la región de la deleción de 5 bp, era una única deleción de 12 bp que había restablecido la pauta de lectura (Figura 17 D). El hecho que en la línea FA104R se hubiese producido una mutación compensatoria de la mutación original en *ERCC4*, que revertía la sensibilidad a al MMC, sugería que efectivamente las mutaciones en *ERCC4* eran la causa de la FA en el paciente FA104.

Después de estos resultados contactamos nuestros colegas y colaboradores de la Universidad de Würzburg en Alemania y analizamos *ERCC4* por secuenciación de Sanger en un total de 18 pacientes alemanes no clasificados y encontramos mutaciones bialélicas en *ERCC4* en otro paciente. El paciente alemán 1333 había nacido en 2002 y se había diagnosticado a los 5 años de edad por una sintomatología clínica caracterizada por malformaciones, BMF y un test de fragilidad cromosómica positivo. De forma parecida al FA104, no se le había detectado ningún rasgo de fotosensibilidad en la piel y las células derivadas del paciente 1333 eran normales para la monoubiquitinación de FANCD2, la formación de *foci* de RAD51 y no presentaban sensibilidad ni a los inhibidores de la Topoisomerasa I ni a los inhibidores de la PARP. La secuenciación de *ERCC4* en el paciente 1333 reveló una duplicación de 28 bp en el exón 11 (c.2371_2398dup28; p.Ile800Thrfs*23; Figura 16 A), que producía una XPF truncada sin el dominio hélice-horquilla-hélice (HhH2) indispensable para la formación del dímero con ERCC1 y su posterior anclaje al DNA (Tripsianes et al., 2005). El otro alelo presentaba una mutación de cambio de sentido que afectaba un aminoácido altamente conservado del dominio helicasa de XPF (c.689T>C; p.Leu230Pro; Figura 18 A y B). El paciente 1333 había heredado del padre la mutación de cambio de sentido mientras que la duplicación de 28 bp era de origen materno. En el análisis de XPF por WB, el paciente 1333 mostraba una banda del tamaño normal, correspondiente al mutante de cambio de sentido, y otra banda correspondiente a una proteína XPF truncada de aproximadamente 90-95 KDa que era compatible con el tamaño predicho para la XPF codificada por el alelo con la duplicación. Las dos bandas se expresaban a niveles muy bajos, alrededor del 5% con respecto al control wt (Figura 18 C) y esta bajada en la expresión de XPF era acompañada también de una bajada en los niveles de la proteína ERCC1. Esto

concordaba con el hecho que las proteínas XPF y ERCC1 se estabilizan mutuamente formando un heterodímero que constituye la endonucleasa activa (de Laat et al., 1998).

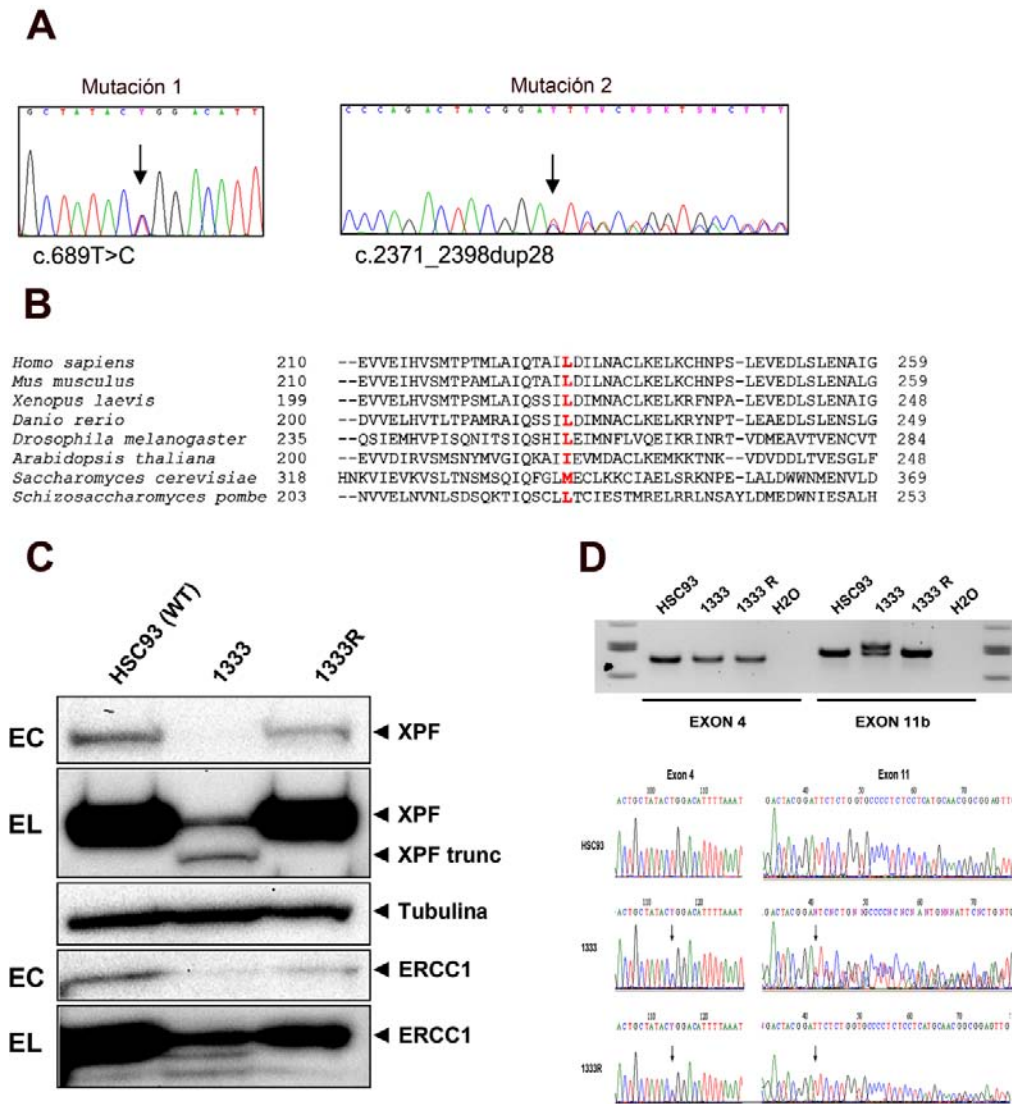


Figura 18: (A) Análisis por secuenciación del DNA del paciente 1333). **(B)** Conservación del residuo L230 en la proteína XPF. **(C)** Análisis por WB que demuestra las dos formas de diferente tamaño (normal y la truncada de 95 KD) y los niveles bajos de la proteína XPF en la línea 1333 y su vuelta a la normalidad en la línea 1333R (EC: Exposición corta; EL: exposición larga). **(D)** La reversión de la duplicación en la línea 1333R: La amplificación por PCR del exón 11 de *ERCC4* en la línea originaria 1333 produce dos bandas: una banda del tamaño esperado y una banda de mayor tamaño por la presencia de la duplicación de 28 bp. La banda de mayor tamaño desaparece en la PCR del exón 8 en la línea 1333R (panel superior). La reversión de la duplicación restablece la secuencia wt (panel inferior) de *ERCC4* en la línea 1333R.

Como en el caso de la línea FA104, se generó una línea 1333 revertida (1333R) resistente a la MMC. En el análisis por WB de la línea 1333R no se encontraba la banda correspondiente a la proteína XPF truncada y, además, los niveles de la banda correspondiente al tamaño normal de la proteína XPF habían vuelto a la normalidad así como los niveles de ERCC1 (Figura 18 C). Se analizó por lo tanto el gen *ERCC4* en la línea 1333R y tanto la amplificación por PCR como la secuenciación del exón 11 revelaron que la duplicación de 28 bp había desaparecido (Figura 18D). Tanto la duplicación como su reversión podrían haber sido provocadas por una repetición invertida de 5 bp que flanquea la región.

Todavía faltaba la prueba formal que efectivamente *ERCC4* era el gen mutado que producía la FA en los dos pacientes y por lo tanto investigamos la complementación genética de la susceptibilidad a la MMC en las LCLs de ambos pacientes mediante la transducción lentiviral del cDNA wt de *ERCC4*. Como se puede apreciar en (Figura 19A y B), solo la transducción del cDNA wt de XPF era capaz de revertir el fenotipo de sensibilidad a los ICL de las líneas FA104 y 1333 mientras que el cDNA de la EGFP, el cDNA wt de FANCA o el cDNA de XPF con la mutación p.Arg689Ser (Figura 19 A y B) no conseguían ningún efecto.

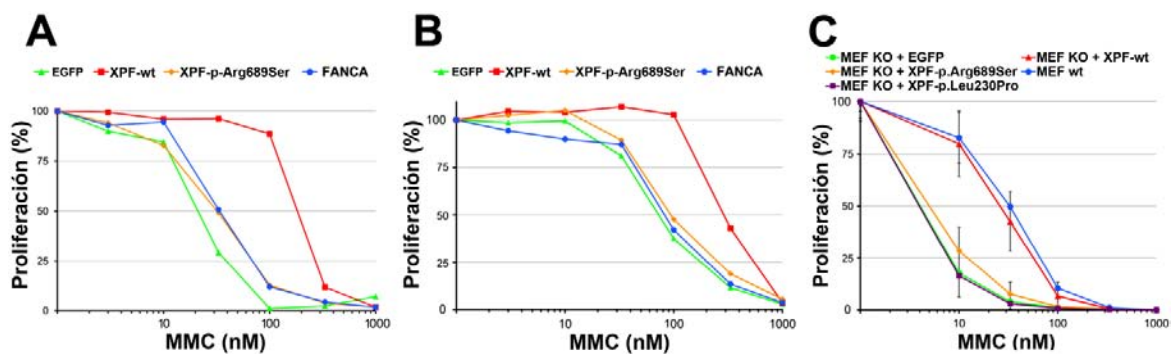


Figura 19: (A) Complementación genética de la sensibilidad a la MMC en la LCLs FA104 y (B) en la LCLs 1333. Las gráficas representan un resultado típico de tres experimentos independientes. (C) Inhibición de la proliferación en MEF *Ercc4*^{-/-} transducidas por medio de vectores lentivirales con la EGFP (control negativo), XPF-wt, XPF-p.Arg689Ser y XPF-p.Leu230Pro). En la gráfica se presentan la media y la desviación estándar de tres experimentos independientes.

La falta de complementación de la sensibilidad a los ICL por parte la proteína XPF-p.Arg689Ser confirmaba que esta mutación era patogénica y que efectivamente perjudicaba la funcionalidad de XPF. La observación que a diferencia de la XPF-wt las proteínas XPF-p.Leu230Pro y XPF-p.Arg689Ser no complementaban la sensibilidad a la MMC de las MEF

Ercc4^{-/-} proporcionó una evidencia más de que ambas mutaciones de cambio de sentido encontradas en los pacientes inactivaban la función de XPF en la ICLR (Figura 17C).

V.3. Mutaciones en *ERCC4* que causan FA afectan marginalmente la NER.

A continuación, tratamos de entender por qué las variantes identificadas llevaban los pacientes a desarrollar FA y no XP o el síndrome de progeria XFE. Nuestra hipótesis fue que las mutaciones que producían la FA en los pacientes resultaban en una fuerte deficiencia en la reparación de los ICL pero estos mutantes tenían que conservar suficiente actividad de la NER como para prevenir la fotosensibilidad en la piel de los pacientes. Para comprobar esta hipótesis, empezamos complementando la línea XP2YO^{SV40}, derivada de un paciente con XPF y caracterizada por una extrema sensibilidad a los UVC, con los mutantes de cambio de sentido XPF- p.Leu230Pro y XPF-p.Arg689Ser y con el XPF-wt. La expresión en la células XP2YO^{SV40} de las proteínas XPF mutantes encontradas en FA conseguía la misma reversión del fenotipo producida por la complementación con el XPF-wt (Figura 20 A). Esto nos confirmaba que los mutantes XPF que causaban FA conservaban su actividad en la NER de las lesiones producidas por UVC. Sucesivamente se investigó la sensibilidad a la radiación UVC de las LCLs FA104 y 1333 y observamos que estas líneas eran menos sensibles con respecto a una LCL de un paciente de XP-A (Figura 20 B). A continuación, estudiamos la sensibilidad a los UVC de la línea 1333^{SV40} comparándola a su contraparte corregida 1333^{SV40} + XPF-wt (Figura 20 C). La 1333^{SV40} tenía una sensibilidad intermedia entre las células que no eran sensibles a los UVC y la XP2YO^{SV40} pero a la dosis más baja se portaba prácticamente como una línea normal (Figura 20 C). Esta observación estaba ulteriormente respaldada por el hecho que los mutantes XPF- p.Leu230Pro y XPF-p.Arg689Ser rescataban aproximadamente el 80% de la sensibilidad a los UVC de las células MEF *Ercc4*^{-/-} (Figura 20 D), mientras que ambos mutantes eran totalmente incapaces de complementar la sensibilidad a la MMC de la misma línea (Figura 19 C). Lamentablemente no se pudieron repetir los mismos experimentos para el paciente FA104 porque había muerto antes de que se lograsen establecer líneas de fibroblastos primarios.

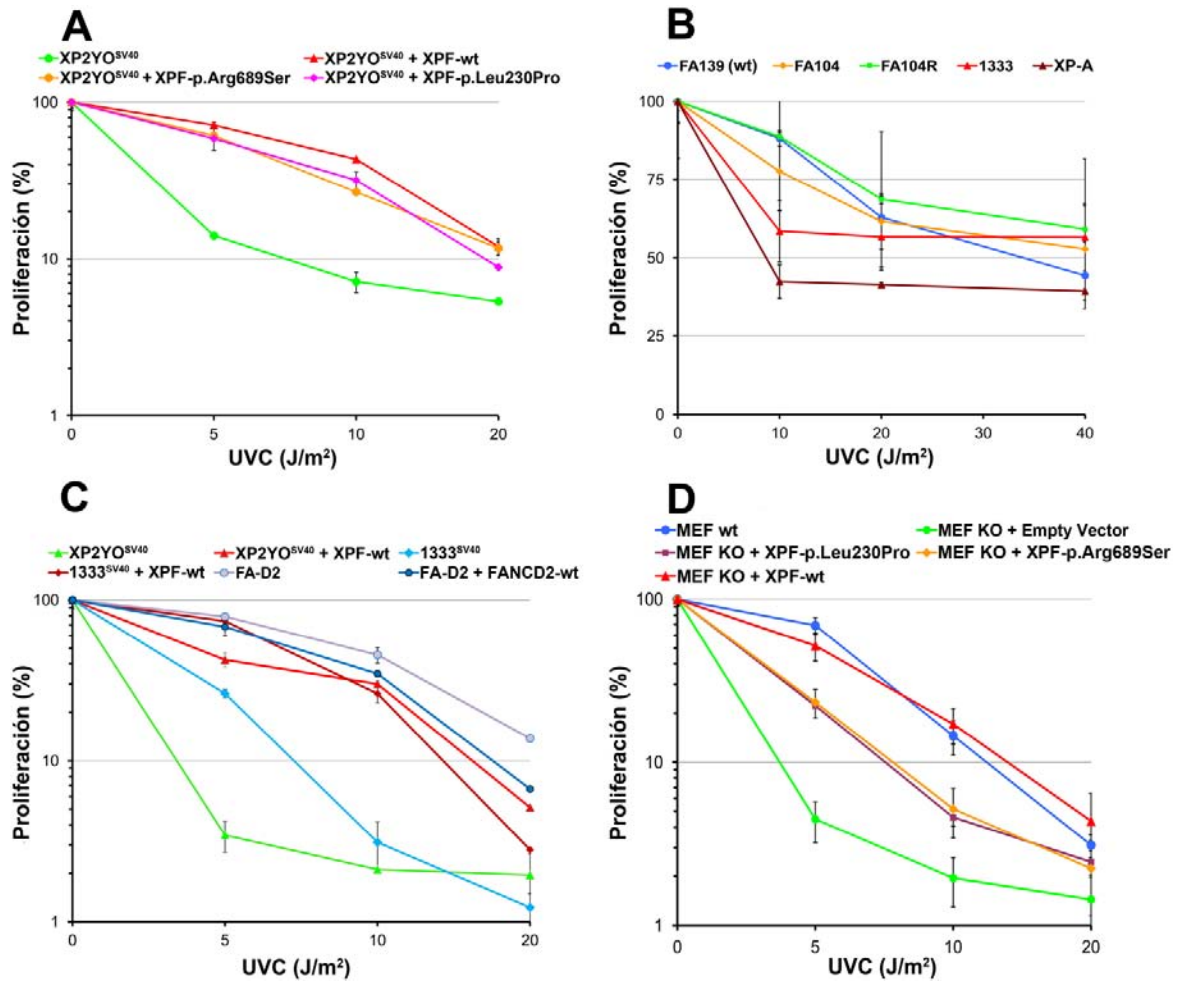


Figura 20: (A) sensibilidad a la radiación UVC de la línea XP-F XP2YO^{SV40} cuando expresa la XPF-wt o los mutantes que producen FA (B) Sensibilidad de las LCLs derivadas de los paciente de FA FA104 y 1333 con respecto a una LCL normal (FA139) o una LCL XP-A. (C) Sensibilidad de las FLC 1333 y 1333 + XPF-wt a la radiación UVC con respecto a los fibroblastos humanos inmortalizados XP-F y FA-D2. (D) La sensibilidad de las MEF *Ercc4*^{-/-} a la radiación UVC puede ser rescatada hasta un 80% por la expresión de los mutantes de XPF que producen FA.

Para investigar más a fondo la NER en el paciente 1333, se midió la UDS inducida por UVC en los 1333^{PF}. Como control se utilizaron los fibroblastos primarios derivados de un paciente XP-F con una sintomatología clínica y una sensibilidad a los UVC muy leve (XP42RO) (Sijbers et al., 1998). Los resultados mostraron que los 1333^{PF} tenían un 24% de UDS, igual o incluso superior a la de los fibroblastos primarios XP42RO (21%) (Figura 21 A y B).

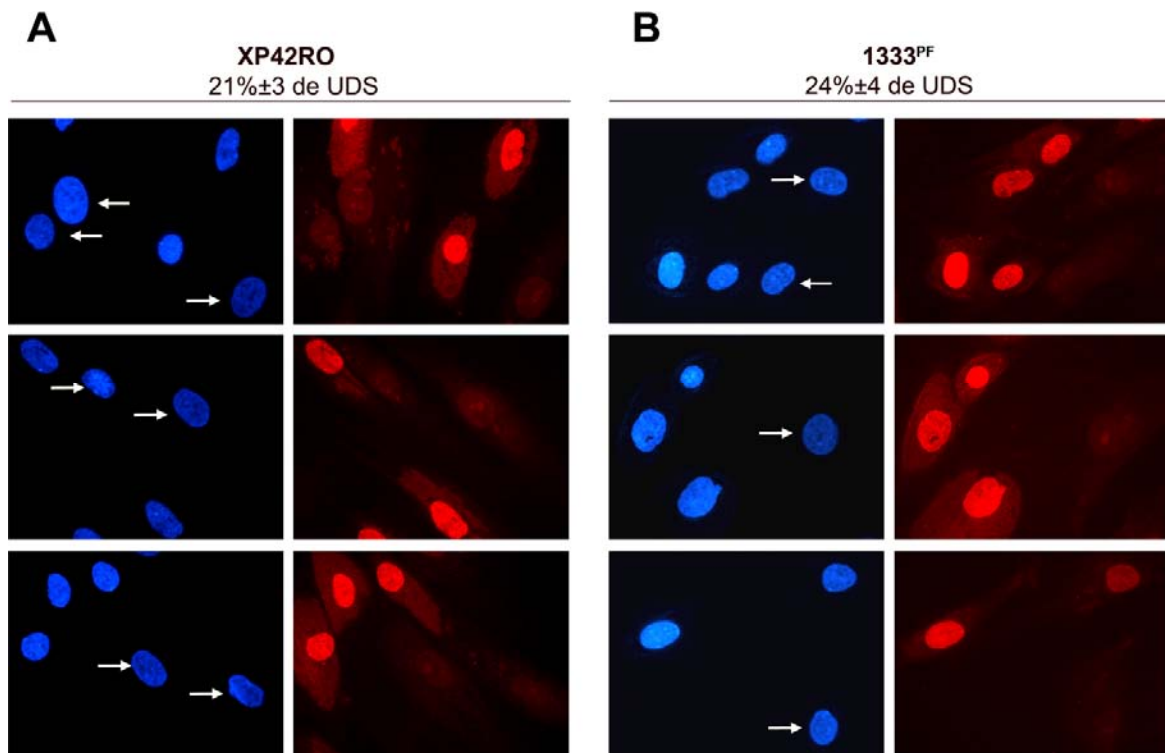


Figura 21: (A) UDS en los fibroblastos primarios XP42RO (XP-F): Los núcleos de las células deficientes por XPF son indicados con las flechas. Se observa el déficit en la síntesis de DNA después de exposición a UVC por la baja intensidad de la fluorescencia roja respecto a las células normales identificables por la presencia de micro esferas de gran tamaño en el citoplasma (véase el apartado Materiales y Métodos, página 61) (B) UDS en los 1333^{PF}.

También se evaluó la UDS en las MEF *Ercc4*^{-/-} que expresaban los mutantes XPF-p.Leu230Pro y XPF-p.Arg689Ser con respecto a los MEF wt: los niveles de la UDS eran del 39% en el caso de las células que expresaban la XPF-p.Leu230Pro y del 44% en el caso de las células que expresaban la XPF-p.Arg689Ser (Figura 22 A). Estos estudios demostraban que los XPF mutantes que causaban FA retenían un nivel de actividad en la NER similar o incluso superior a los de un paciente de XP-F con un fenotipo muy suave. Para apoyar aún más este punto de vista, investigamos la cinética de reparación de los 6-4PP en los sitios de daños inducido por la irradiación localizada con UVC (Volker et al., 2001, Bogliolo et al., 2007) en las líneas derivadas de la XP2YO^{SV40} que expresaban los mutantes XPF-p.Leu230Pro, XPF-p.Arg689Ser y el XPF-p.Ile800Thrfs*23. Las proteínas mutantes de cambio de sentido de XPF eliminaban de forma eficaz los 6-4PP, aunque con menor eficiencia que el XPF-wt, mientras que las células que expresaban la XPF inactivada por la duplicación de 28 bp eran completamente deficientes para la NER (Figura 22 B). Por lo tanto todo demostraba que los alelos de *ERCC4* que presentaban las mutaciones de cambio de

sentido conservaban una actividad residual de la NER más que suficiente para justificar la falta de rasgos de fotosensibilidad en los dos pacientes FA104 y 1333.

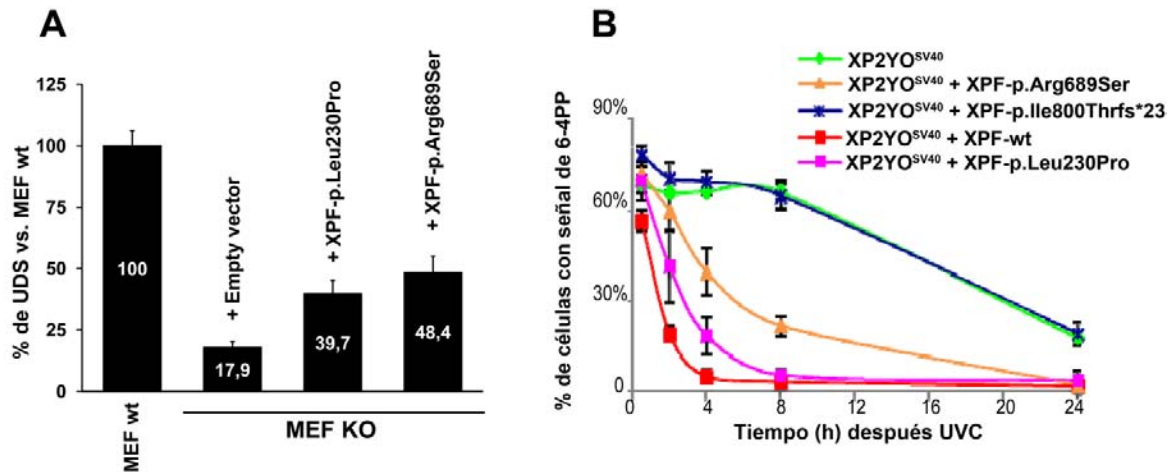


Figura 22 (A) UDS después de irradiación con UVC en las MEF *Ercc4*^{-/-} que expresan los diferentes mutantes de XPF que causan FA. **(B)** Reparación de los 6-4PP en células derivadas de un paciente XPF que expresan los diferentes mutantes FA. Los mutante p. p.Arg689Ser y el p.Leu230Pro reparan eficientemente los daños al DNA inducidos por UVC.

V.4. Mutaciones en *ERCC4* que causan FA inactivan la reparación de los ICL

Decidimos entonces responder a la siguiente pregunta: ¿por qué mutaciones en XPF que causan XP no producen FA o XFE? La respuesta más obvia era que las mutaciones que producían XP no afectaban al papel de XPF en la reparación de los ICL. Por lo tanto evaluamos la sensibilidad de los fibroblastos primarios XP2YO^{PF} (XP-F), XP42RO (XP-F), 1333^{PF} (FA-Q), FA90 (FA-A) y XP51RO (XFE) con el ensayo de fragilidad cromosómica inducida por DEB (Figura 23 A): la línea 1333^{PF} y la línea FA90 dieron, como se esperaba, resultados claramente positivos (Figura 23 A). También la línea XP51RO derivada del paciente XFE resultó tener una fragilidad cromosómica muy acentuada. Por lo contrario, las dos líneas de fibroblastos XP-F XP42RO y XP2YO^{PF} resultaron no tener ninguna fragilidad cromosómica inducida por ICL (Figura 23 A). Los números medios de roturas por célula que se encontraban en las 1333^{PF} (4 roturas) y en la línea XP51RO (8 roturas) estaban en el rango

de valores de fragilidad cromosómica reportada en los linfocitos de sangre de una cohorte de 66 pacientes con FA que se habían estudiado en el nuestro laboratorio utilizando la misma concentración de DEB (desde 1,4 hasta 10 roturas por célula, con una media de 4,3 roturas por célula) (Castella et al., 2011a). Las células 1333 y las XP51RO también resultaron ser las que más se acumulaban en la fase G2/M del ciclo celular después del tratamiento con MMC (Figura 23 B) con un aumento de entre 2 y 2,5 veces de la cantidad de células en G2/M con respecto a las células derivadas de los pacientes XP-F o las células wt. Por lo tanto las células XP, XFE y FA con mutaciones en XPF tenían fenotipos claramente diferentes (Figura 23 A y B) en respuesta a los ICL justificando las diferencias observadas a nivel del fenotipo clínico en los pacientes de las distinta enfermedades.

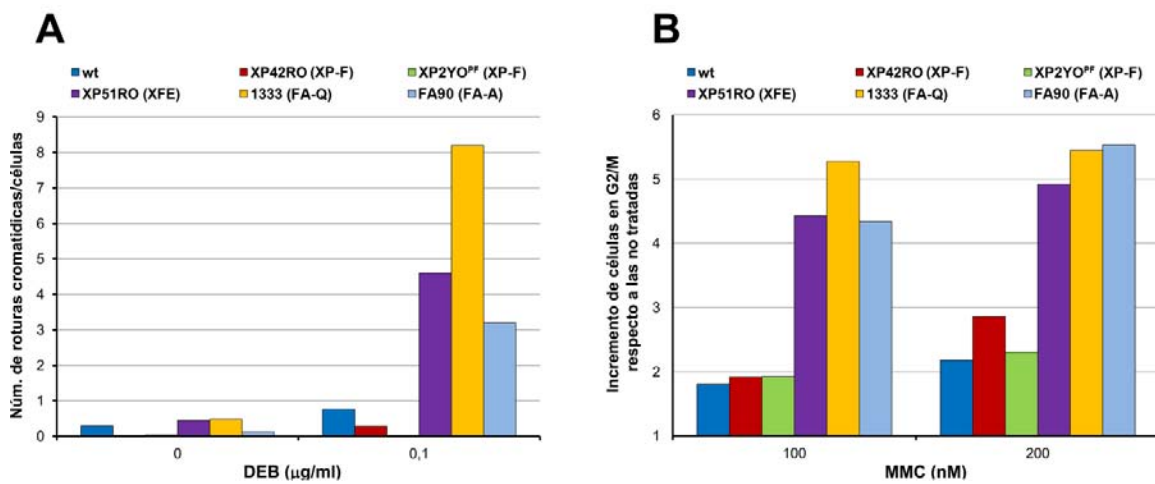


Figura 23: (A) Fragilidad cromosómica inducida por tratamiento con DEB 0,1 µg/ml en fibroblastos primarios. Los fibroblastos primarios con mutaciones en XPF que causan FA (1333) y progeria (XFE; XP51RO) muestran un número de roturas incluso superior a las de un paciente FA-A (FA90). Las células derivadas de pacientes XP-F (XP42RO y XP2YO^{PF}) muestran la misma fragilidad de las células wt. **(B)** De forma coherente con el dato de la fragilidad cromosómica, las células FA con mutaciones en XPF o FANCA se acumulan en G2/M como las células del paciente XFE. Las células XP-F muestran una vez más el mismo fenotipo que las células wt.

V.5. Las proteínas XPF mutadas en FA relocalizan de forma normal al núcleo y al DNA dañado.

Las líneas celulares derivadas de los pacientes XPF muestran un característico fallo de la proteína XPF mutante para trasladarse adecuadamente al núcleo, probablemente debido a un problema de agregación en el citoplasma (Ahmad et al., 2010). Esta característica es aún más evidente en la células del único paciente del síndrome XFE (Niedernhofer et al., 2006). Por lo tanto estudiamos si también las formas mutantes de XPF que causaban FA tenían un defecto en la relocalización al núcleo (Figura 24). Por fraccionamiento celular observamos que las proteínas mutantes XPF que causan FA se podían encontrar en la fracción nuclear (NS) y en la fracción de la cromatina (CR) en cantidades relativas muy parecidas a la XPF de una línea wt (Figura 24 A). Además verificamos por IP que las formas mutadas de XPF seguían interactuando con la proteína ERCC1 de forma parecida al wt (Figura 24 A, panel inferior). También comprobamos por IF que las proteínas XPF que presentaban las mutaciones que causaban FA podían relocalizar normalmente a los sitios de daños en el DNA después de irradiación local del núcleo con UVC (Figura 24 B y C). Esto demostraba que las variantes de XPF que causaban FA se podían trasladar al núcleo y se reclutaban a los sitios donde había NER de forma muy parecida a la proteína XPF normal. Una característica específica de la proteína XPF es su capacidad de interactuar con la proteína SLX4 (FANCP). Analizamos entonces por IP y WB las interacciones entre SLX4 y XPF en las LCLs FA104 y 1333 y encontramos que los mutantes de XPF podían formar complejos con SLX4 de forma parecida a lo que pasaba en las líneas wt utilizadas como control (Figura 24 D). La interacción entre SLX4 y XPF era normal tanto en la línea FA104 como en la 1333 aunque en esta última, debido a los bajos niveles de expresión de XPF y ERCC1, era difícil detectarla a los tiempos de exposición más cortos. Estos resultados sugieren que los mutantes de XPF no causan FA por un defecto en su capacidad de interactuar con SLX4/FANCP.

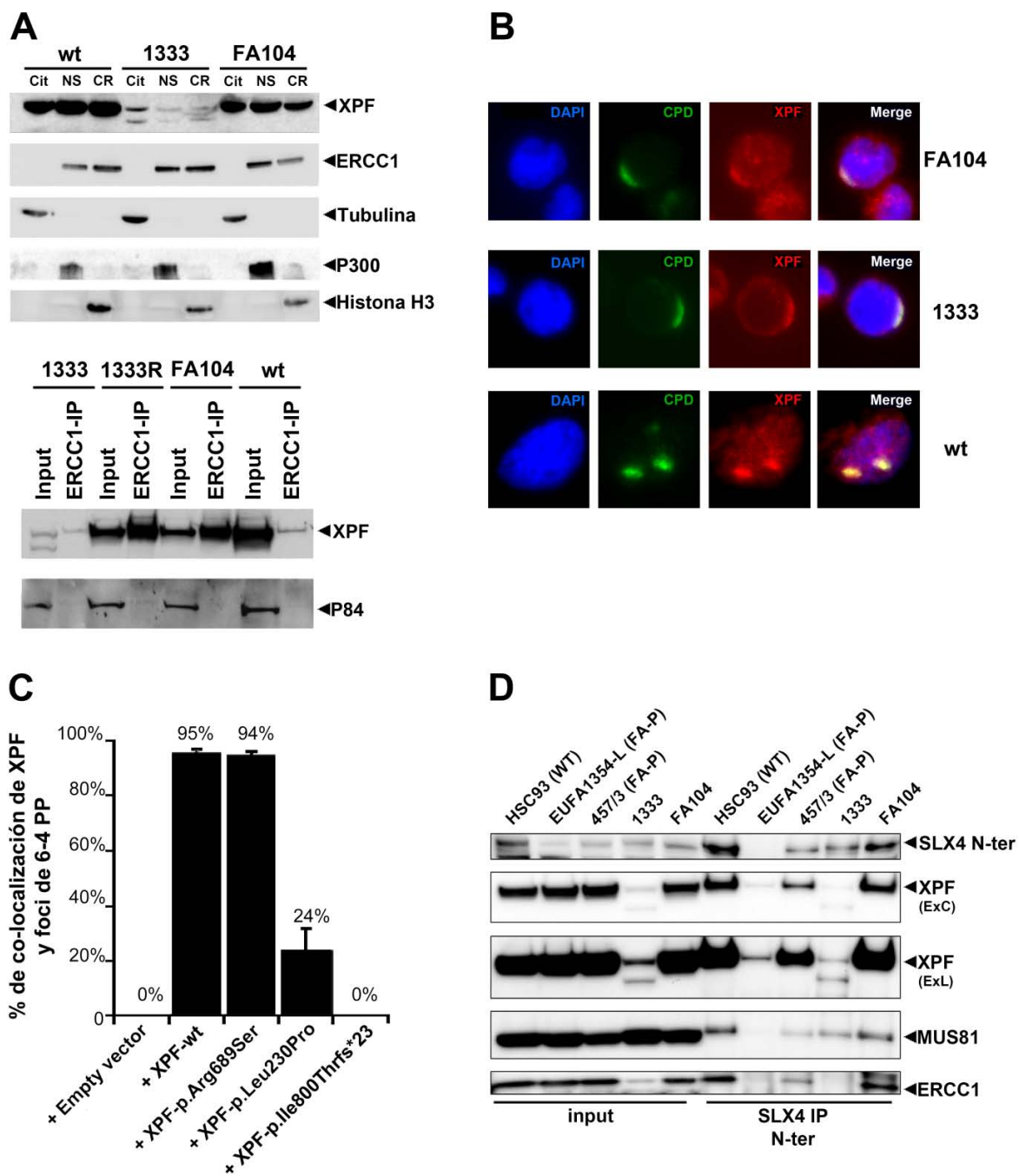


Figura 24: (A) Panel superior: Las proteínas XPF mutantes se encuentran en las diferentes fracciones celulares de forma parecida a la XPF-wt. Controles: fracción citoplasmática (Cit): Tubulina; fracción soluble nuclear (NS): P300; fracción cromatinica (CR): histona H3. Panel inferior: las proteínas XPF mutantes interactúan de forma normal con ERCC1. P84 se utilizó como control de carga para el extracto celular total (Input) (B) Las proteínas XPF mutantes relocalizan de forma normal a los daños inducidos por UVC (CPD) durante la NER y (C) cuantificación de los *foci* de 6-4PP donde se detecta también la presencia de las distintas formas de XPF expresadas en XP2YO^{SV40} (200 *foci* de 6-4 PP, Media +SEM) (D) Las proteínas XPF mutantes interactúan con SLX4 de forma normal: se inmunoprecipitó SLX4 con un anticuerpo contra el N-terminal (SLX4 N-ter) en células wt (HSC93), en células con mutaciones en SLX4 (FA-P) y en las dos líneas FA FA104 y 1333: se detectaron XPF, ERCC1 y MUS81. ExC= exposición corta; ExL=exposición larga.

V.6. Escisión, actividad nucleasa y “unhooking” de los mutantes XPF

A continuación, las proteínas XPF mutantes encontradas en los pacientes con FA se investigaron para averiguar su capacidad de cortar el DNA en proximidad de los ICL mediante dos ensayos *in vitro* (Figura 25 A y 25 B). Para conseguir este objetivo se expresó y purificó el mutante XPF- p.Arg689Ser con el sistema del Baculovirus en células de insecto (Enzlin and Scharer, 2002, Staresincic et al., 2009, Su et al., 2012). Una vez purificadas, se añadieron las proteínas mutantes a extractos de células XP2YO^{SV40} y se realizaron ensayos para estudiar la capacidad de cortar un sustrato de DNA que presenta en una posición concreta una lesión que se repara por NER. Consistente con los datos funcionales anteriores, la proteína purificada XPF- p.Arg689Ser es competente en la etapa de escisión de NER, ya que restaura la capacidad de los extractos celulares de la línea XP2YO^{SV40} de escindir y eliminar un fragmento de DNA dañado de manera similar a la XPF-wt (Figura 25 A).

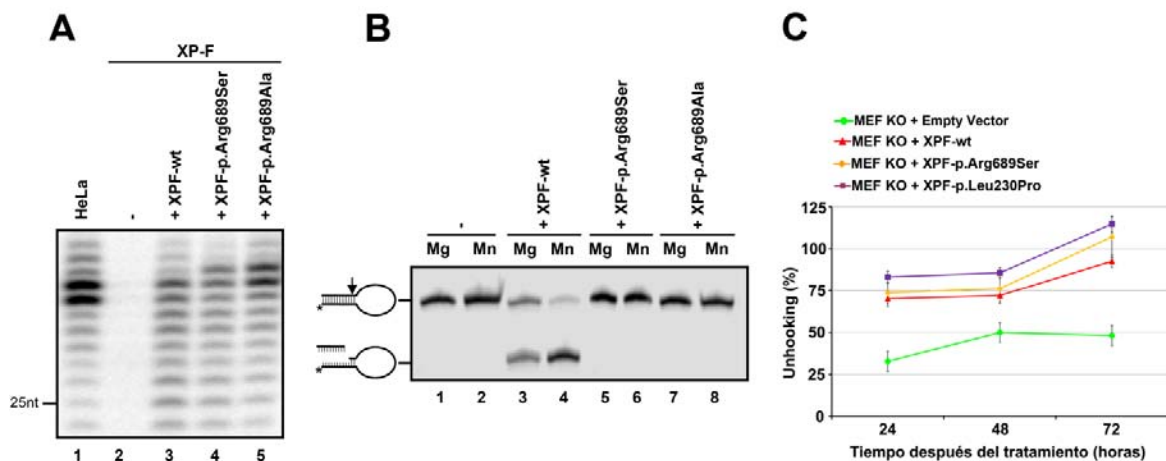


Figura 25: (A) Actividad NER del mutante XPF-p.Arg689Ser. Un plásmido que en un sitio específico contiene una lesión que se repara por NER se incubó con extractos celulares de HeLa o de XP2YO^{SV40} complementados con las diferentes formas purificadas de XPF. En el gel se aprecian los fragmentos del DNA de 24-32 nucleótidos producidos por la NER. (B) la incisión de un sustrato de DNA de conformación Tallo-Bucle por parte de la XPF-p.Arg689Ser. El sustrato se incubó con las formas purificadas de XPF en presencia de 2 mM MgCl₂ o 0,4 mM MnCl₂, y los productos se analizaron por PAGE desnaturalizante. (C) Análisis por COMET de la capacidad de unhooking de MEF *Ercc4*^{-/-} que expresan XPF-wt, XPF-p.Arg689Ser y XPF-p.Leu230Pro después tratamiento con Melfalán. Se midió el Momento de la cola de 200 nucleos y se expresó como porcentaje del Momento de la cola del control no tratado (Unhooking). Las barras de error son SEM.

Sin embargo, la reacción de escisión no es perfecta porque el fragmento escindido es, en promedio, 1 nucleótido más largo de lo esperado (Figura 25 A, carril 4). También se realizó en los ensayos de actividad de nucleasa vitro con XPF- p.Arg689Ser sobre un sustrato de ADN con conformación tallo-bucle (Su et al., 2012). Contrariamente al XPF-wt y a los mutantes encontrados en XP (p.Arg799Trp) o progeria XFE (p.Arg153Pro), el mutante XPF-p.Arg689Ser es incapaz de cortar tal sustrato (Figura 25 B), lo que indica que la actividad nucleasa de XPF- p.Arg689Ser es anormal. Lamentablemente, no se pudieron realizar los mismos experimentos con la proteína mutante XPF- p.Leu230Pro porque tenía muy baja estabilidad y no se pudo purificar en cantidad suficiente para realizar los ensayos bioquímicos. La baja estabilidad de la XPF-p.Leu230Pro a la hora de purificarla concuerda con los bajos niveles de expresión de este mutante que observamos por WB en la LCL 1333 (Figura 20 C). Finalmente estudiamos si las células que expresan las formas mutantes de XPF que producen FA podían cortar el DNA en la etapa de “Unhooking” de la reparación de los ICL. Expresamos de forma estable a los mutantes de XPF en las MEF *Ercc4*^{-/-} y tratamos las células con Melfalán, un medicamento que se emplea en el tratamiento de distintos tipos de cáncer que produce ICL y medimos por COMET la reacción de incisión a varios tiempo después del tratamiento (Figura 25 C) (De Silva et al., 2000). Los dos mutantes (XPF-p.Leu230Pro y XPF- p.Arg689Ser) fueron capaces de complementar el defecto en la etapa de incisión de las MEF *Ercc4*^{-/-} (Figura 25 C), indicando que la sensibilidad de las células FA con mutaciones en XPF podría no ser causada por un defecto en el unhooking de los ICL.

V.7. *ERCC4/FANQ* como posible gen FBOC

Para investigar el posible papel de *ERCC4* en FBOC se cribaron por DHPLC y secuenciación directa los 11 exones codificantes y las regiones intrónicas más cercanas de *ERCC4* en 1.573 casos españoles con alto riesgo de FBOC que habían dado negativo para las mutaciones en los genes *BRCA1/FANCS* y *BRCA2* y en 854 controles sin antecedentes personales o familiares de cáncer. Para más detalles véanse el apartado de Materiales y Métodos, parágrafo VI.1 y (Osorio et al., 2012, Fernandez-Rodriguez et al., 2012, Romero et al., 2011, Bonache et al., 2013). Identificamos así cinco variantes únicas entre los casos y cuatro entre los controles: de las cinco variantes encontradas en los casos, tres eran

potencialmente dañinas, mientras que todas las variantes encontradas en los controles eran supuestamente perjudiciales. Las mutaciones sin sentido se evaluaron utilizando el programa CONDEL (Gonzalez-Perez and Lopez-Bigas, 2011) que predice la patogenicidad de las variantes no sinónimas combinando las herramientas informáticas SIFT, Polyphen2 y Mutation Assessor. Las variantes sin sentido que en el análisis *in silico* resultaron ser potencialmente perjudiciales para la función de XPF se estudiaron posteriormente por análisis funcional. Las variantes supuestamente neutrales se consideraron como tales y no se estudiaron más a fondo (Tabla IX).

Tabla IX: Variantes únicas identificada en el gen *ERCC4* en 1573 casos y 854 controles.

Exón/Intrón	Variante	Cambio en la Proteína	SNVs	MAF	Población	Condel	Ensayo funcional
3	c.448C>T	p.Arg150Cys	rs145402255	–	Casos	Deletérea	Afecta levemente ICLR
3	c.540_541delAG*	p.Arg180Serfs*30	–	–	Controles	–	–
3i	c.584+1G	p.Ile131Serfs*2	–	–	Casos	–	–
5	c.800G>A	p.Arg267His	rs143479220	0,0001	Casos	Neutral	–
9	c.1861A>G	p.Thr621Ala	–	–	Casos	Neutral	–
11	c.2065C>A	p.Arg689Ser	rs149364215	0,0001	Controles	Deletérea	Inhibe ICLR
11	c.2291delG	p.Ser764Thrfs*53	–	–	Controles	–	–
11	c.2357C>T	p.Ser768Phe	rs143081574	–	Casos	Deletérea	Inhibe ICLR

MAF= Frecuencia alélica de la variante en 1000 Genomas

*= Variante encontradas en dos individuos de la población control

En cuanto a los casos, las dos mutaciones de cambio de sentido identificadas, p.Arg150Cys en el exón 3 y p.Ser786Phe en el exón 11 fueron posteriormente investigados funcionalmente. En los controles identificamos dos mutaciones de frameshift c.540_541delAG en el exón 3 (esta variante se encontró de forma independiente en dos individuos) y la variante c.2291delG en el exón 11. La mutación c.2291delG produce una proteína XPF truncada sin el dominio HhH2 necesario para la dimerización con ERCC1 (de Laat et al., 1998), de una forma muy parecida a la mutación c.2371_2398dup28; p.Ile800Thrfs*24 que compromete ambas funciones de XPF en la NER y la ICLR, como se describe anteriormente. La última variante dañina encontrada en los controles era la mutación de cambio de sentido p.Arg689Ser previamente caracterizada en el paciente FA FA104 que vimos afectar al papel de XPF en la ICLR (Bogliolo et al., 2013). Para evaluar su impacto funcional, las variantes de cambio de sentido c.448C>T, c.2065C>A y c.2357C>T se

introdujeron por mutagénesis dirigida en un el cDNA wt de *ERCC4* y las construcciones resultantes fueron transducidas en MEF *Ercc4*^{-/-} (Figura 26).

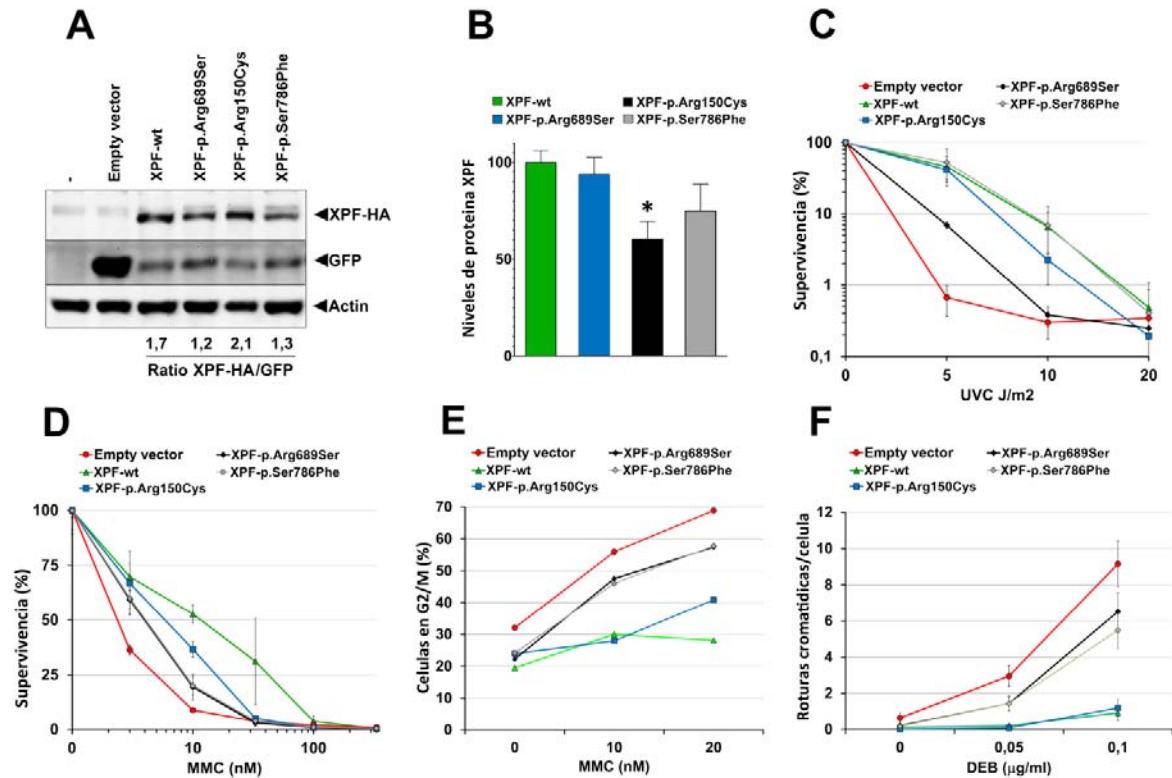


Figura 26: Estudios funcionales en MEF *Ercc4*^{-/-} de las variantes de cambio de sentido de XPF. **(A)** análisis por WB de las expresiones de HA-XPF y de la GFP en los MEF *Ercc4*^{-/-} transducidos con las diferentes formas de XPF. **(B)** Los niveles de XPF se cuantificaron y se normalizaron a los niveles de la GFP. Los niveles normalizados de XPF se muestran como porcentaje de los niveles de la proteína XPF-wt. En la gráfica se presentan la media y SEM de 6 experimentos (* p = 0,005 prueba de t-student). **(C)** Sensibilidad a los UVC de los MEF *Ercc4*^{-/-} que expresan los diferentes formas de XPF. Las células se analizaron 72 h después de la irradiación. Los datos representan medias y SD de tres experimentos independientes. **(D)** Sensibilidad a la MMC en las mismas células que (C). Los datos representan medias y SD de dos experimentos independientes. **(E)** Parada en la fase del G2/M del ciclo celular por MMC en las mismas células que en (C). **(F)** Fragilidad cromosómica inducida por DEB en las mismas células que en (C). Se muestran en la gráfica el número medio (± SEM) de roturas cromatídicas por célula en observadas en 25 metafase por cada concentración de DEB.

La variante p.Arg689Ser reduce aproximadamente un 40% la estabilidad de XPF, mientras que las variantes p.Arg150Cys y p.Ser786Phe no tienen impacto alguno sobre de la estabilidad de la proteína (Figura 26 A y B). La expresión tanto de la variante p.Arg150Cys como de la variante p.Ser786Phe complementa la sensibilidad de las MEF *Ercc4*^{-/-} a los UVC (Figura 26 C) mientras que los MEF que expresan la variante p.Ser786Phe o la

p.Arg689Ser presentan un fenotipo FA de sensibilidad a la MMC, de parada en la fase G2/M del ciclo celular inducida por ICL y de fragilidad cromosómica (Figura 26 D, E y F). Estos datos confirman que, asemejándose a la p.Arg689Ser, la variante p.Ser786Phe impide específicamente la reparación de los ICL y, por lo tanto, podría ser una mutación capaz de producir FA. Notablemente ambas mutaciones se encuentran en el dominio nucleasa de XPF. A pesar de una sensibilidad leve a la MMC (Figura 26 D y E), las MEF *Erc4*^{-/-} que expresan la variante p.Arg150Cys no muestran los otros rasgos del fenotipo FA (Figura 26 F). Estos datos indican un impacto casi nulo de la variante c.448C>T sobre de las funciones de XPF.

DISCUSIÓN

“...In the end nature decides what is relevant for human disease.”

Johan P. de Winter

VI.DISCUSIÓN

VI.1. Caracterización genética mediante WES + MLPA de 49 pacientes FA.

La elevada heterogeneidad genética de la FA es uno de los principales problemas para el diagnóstico molecular de la enfermedad. En los últimos años, para optimizar la detección de mutaciones, se han desarrollado diversas estrategias basadas principalmente en la complementación y la caracterización por secuenciación tradicional del DNA de los genes FA (Casado et al., 2007, Castella et al., 2011b, Ameziane et al., 2008). Con el fin de mejorar y acelerar este proceso, evaluamos la posibilidad de aplicar la WES acoplada con estrategias de MLPA para identificar tanto las mutaciones puntuales como las grandes deleciones intragénicas que son relativamente comunes en FA (Morgan et al., 1999). En la población española, cerca del 80% de los pacientes pertenecen al grupo de complementación FA-A y la gran prevalencia de este grupo es típica de las poblaciones mediterráneas (Savoia et al., 1996, Bouchlaka et al., 2003, Savino et al., 1997) (Figura 27 A). La subpoblación de 45 pacientes españoles (4 pacientes de nuestra cohorte de 49 son originarios de Reino Unido) analizada en nuestro estudio tiene pero una distribución diferente respecto a la población española de pacientes FA (Figura 27 B). Esta discrepancia es debida principalmente a que nuestra cohorte no es completamente al azar: la mayoría de los pacientes que se incluyeron en la población estudiada son pacientes que hasta ahora no se habían podido caracterizar por completo utilizando los métodos tradicionales de subtipaje y de caracterización molecular. Estos métodos están optimizados para los subtipos generalmente considerados más comunes (FA-A, FA-C y FA-G) (Auerbach, 2009) y por este motivo hay una mayor presencia en nuestra subpoblación de subtipos raros (FA-D2, FA-G, FA-D1) de FA (Figura 27 B).

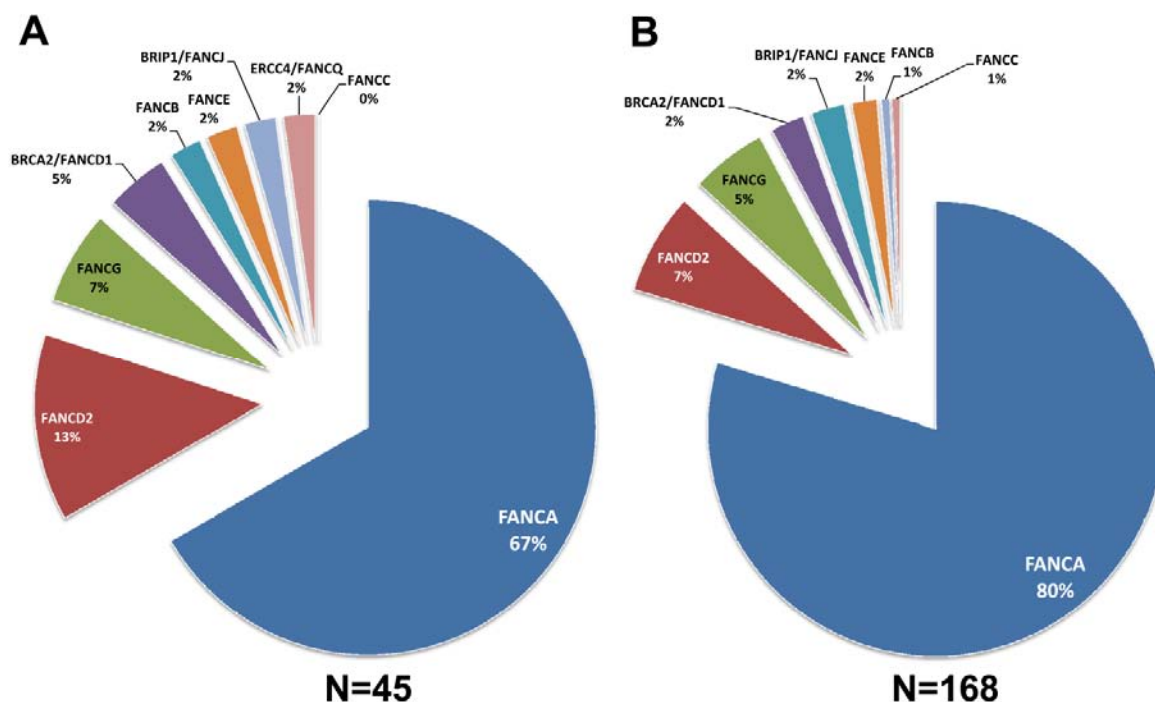


Figura 27: (A) Distribución de los grupos en la población de pacientes españoles analizados según el método WES + MLPA y (B) Distribución de los diferentes grupos de complementación en la totalidad de la población de pacientes de FA españoles

Solamente 9 de los pacientes incluidos en nuestra cohorte estaban totalmente para estudiar. Pacientes como el FA049, FA050 o FA090, por ejemplo, entraron en nuestra base de datos en 2001 y aun no se había conseguido caracterizarlos por completo. El paciente FA158 por ejemplo, entró en nuestro Biobanco en el 2003 y desde entonces no fue posible ni tan solo determinar su grupo de complementación. Con nuestro análisis hemos determinado que el paciente tenía mutaciones bialélicas en *BRCA2/FANCD1*, un gen que, debido a su gran tamaño, es muy complicado de analizar con los métodos tradicionales. Por lo tanto, nuestra población estaba enriquecida por casos complicados (los raros de los raros) que, en algunos casos, habían estado casi 15 años en nuestro Biobanco sin poderse caracterizar por completo debido a los límites de los métodos tradicionales. No obstante, fuimos capaces de conseguir caracterizar por completo el 85,7% de los pacientes. Como reportado en la Tabla VI los pacientes FA166, FA388 y FA121 encontramos solo una mutación en *FANCA* (FA166 y FA388) y *FANCE* (FA166) pero el gen afectado por esa mutación es concorde con el gen encontrado con el análisis por subtipaje retroviral. En el caso de los pacientes FA664 y FA681 (hermanos y no previamente subtipados) encontramos una mutación en *FANCD2* y

los fibroblastos primarios del paciente FA681 no presentan banda específica de FANCD2 en el análisis por WB (no mostrado). Los pacientes FA572 y FA124 presentan una mutación claramente patogénica en *FANCA* y *FANCD2* respectivamente. *FANCA* y *FANCD2* los genes más frecuentemente mutados en la población española (Figura 27) y no tenemos ninguna otra evidencia por WES de mutaciones en otros genes que podrían explicar el fenotipo claramente FA de los dos pacientes. Después de nuestro análisis ya no existen pacientes sin subtipo determinado en nuestra base de datos.

Nuestra cohorte de pacientes estaba constituida por 34 pacientes Full-FA y 15 pacientes mosaicos. En el caso de los pacientes Full-FA fuimos capaces de detectar las dos mutaciones en 31/34 pacientes, caracterizando así por completo el 91,2% de los pacientes, mientras que conseguimos caracterizar por completo solo el 73% de los individuos mosaicos (11/15). Si se restringe el análisis solo a las muestras de PB, se evidencia como la eficiencia de nuestro método baja hasta un 60% (6/10) a la hora de caracterizar por completo los pacientes mosaicos. En el caso de los Full-FA fuimos capaces de detectar las dos mutaciones en 27 de las 29 muestras de DNA proveniente de PB (93,1%). Esta diferencia ($p < 0.05$, chi-squared test) se debe probablemente al efecto del mosaicismo somático que puede haber eliminado una de las mutaciones a nivel de la médula y, por tanto, también a nivel de la PB (Soulier et al., 2005). En base a esta observación y a otras ya reportadas en la literatura, se deduce que también para el diagnóstico molecular utilizando el método WES+MLPA en el caso de los pacientes mosaicos la fuente de DNA ideal son los fibroblastos de piel (Alter, 2014, Alter et al., 2005).

Este estudio ha permitido la elaboración de un método de trabajo que tiene ventajas evidentes respecto al análisis tradicional. El análisis por WES + MLPA elimina por completo la necesidad del subtipaje retroviral, una aproximación que no se puede realizar fuera de los centros de investigación ultra especializados por sus costes elevados y la necesidad de personal altamente cualificado. Utilizando nuestra estrategia se obtienen resultados en un tiempo extremadamente inferior al subtipaje tradicional (alrededor de un mes desde el envío del DNA y la recepción de los datos de la WES) y también son más fiables. Como ya presentamos en la sección “Resultados” de esta tesis, en nuestra población había 10 pacientes (acerca del 20% del total de la población analizada) en que se habían descartado varios genes pero, una vez sometidos al análisis por WES, 7 de estos se revelaron ser falsos negativos por

tener mutaciones en genes que se habían descartado por subtipaje retroviral. Otro aspecto a tener en cuenta de nuestra estrategia es que la MLPA, una aproximación muy costosa, se aplica posteriormente a la WES y solo en una subpoblación reducida de pacientes FA donde no se han encontrado las dos mutaciones. En el método tradicional, el primer paso para la caracterización de las mutaciones en todos los pacientes FA-A es un análisis por MLPA (Castella et al., 2011b). En nuestro caso tuvimos que recurrir a la MLPA solo en 10 de los 34 pacientes FA-A analizados, reduciendo así más del 70% la utilización del método de MLPA. Nuestro trabajo es el primero por número de pacientes FA analizados por WES y por utilizar en nuestro análisis exclusivamente un kit de captura comercial para el exoma entero. Otros estudios ya se habían publicado pero con un número limitado de pacientes o utilizando paneles de genes (Zhang et al., 2015, Knies et al., 2012, Schuster et al., 2013, Ameziane et al., 2012) o acoplado la *Whole transcriptome shotgun sequencing* (secuenciación del transcriptoma o RNA-seq) y CGH arrays a la WES para lograr la identificación de todas las posibles mutaciones (Chandrasekharappa et al., 2013). En nuestro caso lo que se quería lograr era un método que fuese cuanto más rápido, sencillo y económico posible para la caracterización de los pacientes FA y que todavía mantuviese una eficacia suficientemente alta como para poderse utilizar de forma rutinaria en lugar del método estándar.

Entre todas las mutaciones nuevas encontradas en nuestra cohorte, 5 de ellas son mutaciones de cambio de sentido (2 en *FANCA*, 1 en *FANCB*, 1 en *BRCA2/FANCD1* y 1 en *FANCO/ERCC4*). La patogenicidad de estas mutaciones ha sido validada mediante análisis *in silico* utilizando distintos algoritmos de predicción de efecto de la variante. En todos los casos los dos algoritmos aplicados, teniendo en cuenta el análisis de la diferencia de propiedades entre los aminoácidos y la conservación de la posición en la evolución, coinciden en que el cambio tiene consecuencias para la funcionalidad de la proteína. La mutación de cambio de sentido c.2065C>A (p.Arg689Ser) en el gen *FANCO/ERCC4* se confirmó ser perjudicial para las funciones de la proteína XPF en la reparación de los ICL mediante ensayos funcionales (Figuras 17 y 26) (Bogliolo et al., 2013, Osorio et al., 2013). Para confirmar la patogenicidad de las restantes variantes estamos desarrollando ensayos funcionales *ad hoc* pero, de momento, no tenemos confirmaciones empíricas de los resultados del análisis bioinformático.

Una alternativa a la WES en el diagnóstico de la FA, es la NGS de paneles de genes. Se pueden desarrollar kits de captura únicamente para las secuencias exónicas de los genes involucrados en la FA evitando así de tener que secuenciar todo el exoma. Los paneles de genes pueden ser una alternativa muy válida para el diagnóstico molecular de la FA: teóricamente es posible llegar a tener coberturas mucho más altas respecto a la WES tradicional y esto permite la detección más fiable de las variantes y también la detección de las deleciones intragénicas a través de las variaciones en la cobertura de las secuencias delecionadas (Knies et al., 2012, Ameziane et al., 2012). Los paneles de genes además pueden ser informativos a nivel de la caracterización de los pacientes mosaicos porque pueden permitir la detección de las mutaciones también cuando son presentes en pocas células a nivel de la PB (Jamuar et al., 2014). Pero la gran ventaja que la WES tiene respecto a la NGS de paneles de genes, es que la WES permite acoplar el diagnóstico con la investigación de base. De hecho únicamente la WES puede permitir el descubrimiento de variantes en genes hasta al momento no asociados a una enfermedad hereditaria y, además, en un futuro podría permitir la identificación de variantes génicas que, sin ser la causa directa, podrían modular la severidad de la enfermedad.

Analizamos las deleciones intragénicas que encontramos mediante el análisis por MLPA también con el software IGV y todas las deleciones se podían detectar utilizando la variación de los datos de cobertura (Figura 14). Esta evidencia nos hizo decidir de empezar a desarrollar en colaboración con el Laboratorio de Genómica computacional del Centro de Investigación Príncipe Felipe (CIPF) de Valencia, un *pipeline* bioinformático para el análisis estadístico de la variación de la cobertura con el fin de poder determinar correctamente la presencia de deleciones intragénicas en hetero- y homocigosis en todos los genes FA. Por lo tanto, en un futuro cercano, el análisis por WES podría ser suficiente para el análisis molecular completo de los pacientes FA sin la necesidad de recurrir al análisis por MLPA. Además utilizando la WES se podrán detectar las deleciones intragénicas en todos los genes FA y no solamente en *FANCA*.

VI.1.1. Espectro mutacional de los pacientes FA-A analizados

El espectro mutacional de nuestra cohorte de pacientes FA-A (Tabla X) revela, como era esperado, una heterogeneidad de mutaciones muy elevada. En los 34 pacientes FA-A analizados se detectaron 37 mutaciones distintas en el gen *FANCA* y 12 de estas (sin contar las grandes deleciones) no habían sido descritas anteriormente en ninguna otra población. El tipo de mutación más frecuente en nuestra población son las InDel (25 alelos de 65) que se detectan en 17 de 34 de nuestros pacientes (50%). Con la sola excepción de la deleción 3788_3790delTCT (Magdalena et al., 2005) que produce la pérdida completa de un codón (p.Phe1263del) manteniendo pero la pauta de lectura de *FANCA*, todas las otras InDel dan lugar a un desplazamiento del marco de lectura que produce una interrupción prematura de la traducción. La mutación más frecuentemente encontrada en nuestra población de pacientes FA-A es la deleción c.1115_1118delTTGG (p.Val372Alafs*42) (Fanconi anaemia/Breast cancer, 1996) con una prevalencia alélica de 8/65 y presente en 7 pacientes de 34 (20,6%). Las otras 2 mutaciones más frecuentemente encontradas en los pacientes FA-A de la población española la c.3788_3790delTCT y la c.3558insG en nuestra cohorte tienen ambas una prevalencia alélica de 5/68, ambas presentes en 4 pacientes. El 23,5% (8/34) de los pacientes caracterizados tienen por lo menos una mutación de cambio de sentido para un total de 6 mutaciones distintas y de 8 alelos de los 65 en el conjunto de los pacientes FA-A. Las dos mutaciones de cambio de sentido no reportadas se examinaron con las herramientas informáticas SIFT (Kumar et al., 2009) y Polyphen-2 (Adzhubei et al., 2010) y en ambos análisis resultaron tener una elevada probabilidad de ser patogénicas (Tabla VII). No había diferencia significativa entre los pacientes españoles presentes en nuestra subpoblación y el total de la población FA-A española en la prevalencia de los tipos de mutaciones en *FANCA* (Figura 28). Las deleciones intragénicas son muy comunes en el gen *FANCA* y se sabe que el origen de las deleciones se debe, en la mayoría de los casos, a recombinaciones entre elementos *Alu* colocados en *cis* y pérdida del fragmento interno (Callen et al., 2004, Morgan et al., 1999, Castella et al., 2011b). En el 26,5% (9/34) de los pacientes FA-A de nuestra población se han detectado 8 distintas deleciones intragénicas en el gen *FANCA*.

Tabla X: espectro mutacional de los pacientes FA-A analizados.

Mutaciones sin sentido: 20%					
Cambio en DNA	Exón	Proteína	Referencia	Alelos	Pacientes
c.65G>A	1	p.Trp22*	Savoia , 1996	2	1
c.163C>T	2	p.Gln55*	Wijker , 1999	2	1
c.4130C>G	41	p.Ser1377*	Ameziane , 2008	2	2
c.295C>T	4	p.Gln99*	Callen , 2004	1	1
c.790C>T	8	p.Gln264*	Savino , 1997	1	1
c.1033G>T	12	p.Glu345*	No reportada	1	1
c.2128A>T	23	p.Arg710*	No reportada	1	1
c.2641C>T	28	p.Gln881*	Castella , 2011b	1	1
c.3763G>T	37	p.Glu1255*	No reportada	2	1
Mutaciones de cambio de sentido: 12,3%					
Cambio en DNA	Exón	Proteína	Referencia	Alelos	Pacientes
c.3263C>T	33	p.Ser1088Phe	Rockfeller DB	2	2
c.3913C>T	39	p.Leu1305Phe	Rockfeller DB	2	2
c.1100G>T	13	p.Ser367Ile	No reportada	1	1
c.2574C>G	27	p.Ser858Arg	Tamary , 2000	1	1
c.3335T>G	33	p.Val1112Gly	No reportada	1	1
c.3982A>G	40	p.Thr1328Ala	Rockfeller DB	1	1
Inserciones/delecciones: 38,5%					
Cambio en DNA	Exón	Proteína	Referencia	Alelos	Pacientes
c.233_236delTTGA	3	p.Thr78Ilefs*16	Callen , 2004	1	1
c.472insC	5	p.Hist158Profs*23	No reportada	1	1
c.1860dupC	21	p.Tyr621Leufs*9	No reportada	1	1
c.3448dupC	35	p.Leu1150Phefs*65	No reportada	1	1
c.4124_4125delCA	41	p.Thr1375Serfs*49	Rockfeller DB	1	1
c.2837insT	29	p.Ser947Phefs4*	No reportada	2	1 †
c.3558insG	36	p.Arg1186Glufs*28;	Savino, 2003	5	4
3788_3790delTCT‡	38	p.Phe1263del;	Magdalena , 2005	5	4
c.1115_1118delTTGG	13	p.Val372Alafs*42	Fanconi anaemia/Breast cancer, 1996	8	7
Splicing: 9,2%					
Cambio en DNA	Exón	Proteína	Referencia	Alelos	Pacientes
c.709+5G>A	7i	p.D237_F238delinsGAFMTRCGFLD	Wijker , 1999	1	1
c.893+2T>C	10i	ND	no reportada	1	1
c.893+1G>T	10i	ND	Savino , 2003	1	1
c.3828+1G>T	38i	ND	no reportada	1	1
c.1083+2T>C	12i	ND	no reportada	2	2
Delecciones intagénicas: 20%					
Delección	Alelos	Pacientes			
Exón 3	2	1			
Exones 1-6	2	1 †			
Exones 16-20	2	1 †			
Exones 18-21	2	1 †			
Exones 21-29	2	2			
Exones 15-17	1	1			
Exones 27-43	1	1			
Exones 32-33	1	1			

‡: 3788_3790delTCT produce la pérdida de un aminoácido pero conserva la pauta de lectura

†: Alelos encontrados en individuos originarios de Reino Unido

Finalmente, las mutaciones que afectan el *splicing* son las menos frecuentes y solo 6/34 (17,6%) de los pacientes FA-A analizados presentan mutaciones de *splicing*. Tres de las mutaciones de *splicing* encontradas no se habían nunca reportado antes, todas pero afectan a la secuencia consenso GU del extremo 5' del intrón (*donor site*) y por lo tanto se consideran deletéreas sobre del correcto procesamiento del RNA primario y sobre de la traducción de la

proteína (Baralle and Baralle, 2005). No hay diferencia significativa entre la prevalencia de los distintos tipos de mutaciones en *FANCA* de nuestra subpoblación y la de la población española entera de FA-A.

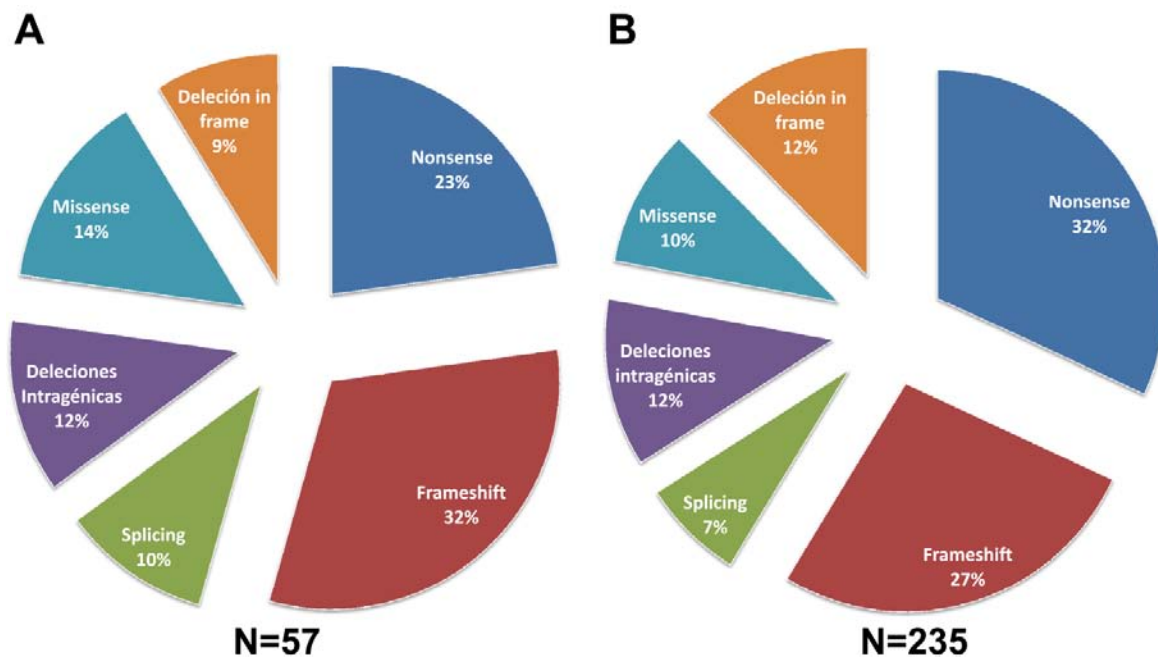


Figura 28: (A) Prevalencia alélica de los distintos tipo de mutaciones en *FANCA* en la población de pacientes españoles analizados según el método WES + MLPA y (B) en la totalidad de la población de pacientes FA-A españoles

VI.1.2. Espectro mutacional de los pacientes FA-A españoles

Los pacientes FA-A constituyen el 80% de la población española de FA (Figura 27). Actualmente en nuestra base de datos hay registrados 133 individuos FA-A en total, 121 de los cuales se han caracterizado molecularmente. Hasta el día de hoy se han encontrado 235 alelos mutados sobre de un total de 242 alelos mutados de *FANCA* (97%) y la prevalencia alélica de los distintos tipos de mutaciones se puede observar en Figura 28, Figura 29 y en la Tabla XI. Después de la mutación “gitana” (c.295C>T; p.Gln99*(Callen et al., 2005)) con 52 alelos, la mutación más prevalente en la población española (y también en el mundo (Levrán

et al., 2005b, Magdalena et al., 2005)) es la c.3788_3790delTCT; p.Phe1263del (29 alelos). Esta mutación produce la pérdida de un aminoácido sin perturbar al marco de lectura del mRNA. En líneas celulares derivadas de pacientes homocigotos para la mutación c.3788_3790delTCT los niveles de FANCA_{p.Phe1263del} son muy parecidos a los de la FANCA-wt, pero la FANCA-p.Phe1263del no puede entrar en el núcleo. Esta incapacidad de trasladarse al núcleo es una característica también de otros mutantes FANCA de cambio de sentido (Castella et al., 2011b). En la población española de pacientes FA-A hay un total de 37 pacientes con una mutación de cambio de sentido o la mutación c.3788_3790delTCT (37/121; 30,6%) mientras que el resto de los pacientes presenta mutaciones bialélicas que potencialmente afectan a la integridad de FANCA. Esto es importante de cara al desarrollo de nuevas terapias o nuevos fármacos para tratar la enfermedad: se puede suponer que una terapia farmacológica para reactivar la ruta FA/BRCA2 en los pacientes FA-A tendría mayor probabilidad de éxito si las mutaciones no comprometen del todo la estabilidad de FANCA. Ya se ha demostrado que la sobreexpresión de determinados mutantes de cambio de sentido de FANCA es capaz de reactivar la ruta FA/BRCA2 en células nulas para *FANCA* (Adachi et al., 2002). En el caso de mutaciones bialélicas que trunquen o desestabilicen por completo FANCA la terapia génica podría ser la única alternativa de cura si no fuese posible encontrar un fármaco en grado de reactivar la ruta (monoubiquitinación de FANCD2 o formación de *foci* de FANCD2) en ausencia de *FANCA*.

En la literatura suele reportarse que un 15–20% de los pacientes *FANCA* presentan grandes deleciones intragénicas (Ameziane et al., 2008, Gille et al., 2012) este porcentaje puede llegar hasta el 40% dependiendo de la población analizada (Morgan et al., 1999). En línea con cuanto ya publicado, la población española de FA-A comprende un 20% de pacientes (25/121) que presentan por los menos una deleción intragénica. En la Figura 29 podemos observar la distribución a lo largo de *FANCA* de todas las mutaciones identificadas hasta el día de hoy en la población española de FA.

Tabla XI: Espectro mutacional de los pacientes FA-A españoles

Mutaciones sin sentido: 31,9%					
Cambio en DNA	Exón	Proteína	Referencia	Alelos	Pacientes
c.295C>T	4	p.Gln99*	Callén, 2005	52	27
c.65G>A	1	p.Trp22*	Rockfeller's FA Database	4	3
c.4130C>G	41	p.Ser1377*	Ameziane, 2008	4	4
c.1771C>T	19	p.Arg591*	Morgan, 1999	3	2
c.163C>T	2	p.Gln55*	Wijker, 1999	2	1
c.2641C>T	28	p.Gln881*	Castella, 2011b	2	2
c.258T>A	3	p.Tyr86*	Rockfeller's FA Database	1	1
c.718C>T	8	p.Gln240*	Gille, 2012	1	1
c.790C>T	8	p.Gln263*	No reportada	1	1
c.1033G>T	12	p.Glu345*	No reportada	1	1
c.1147delG	13	p.Glu383Argfs*32	Castella, 2011b	1	1
c.1874G>C	21	p.Cys625Ser	Castella, 2011b	1	1
c.2128A>T	23	p.Arg710*	No reportada	1	1
c.3762G>T	37	p.Glu1255*	No reportada	1	1
Mutaciones de cambio de sentido: 9,8%					
Cambio en DNA	Exón	Proteína	Referencia	Alelos	Pacientes
c.3913C>T	39	p.Leu1305Phe	Rockfeller's FA Database	5	5
c.1475A>G	16	p.His492Arg	Adachi, 2002	2	1
c.2303T>C	26	p.Leu768Pro	Gille, 2012	2	1
c.3263C>T	33	p.Ser1088Phe	Wijker, 1999	2	2
c.4198C>T	42	p.Arg1400Cys	Rockfeller's FA Database	2	2
c.821C>T	9	p.Leu274Pro	Castella, 2011b	1	1
c.1100G>T	13	p.Ser367Ile	No reportada	1	1
c.2224C>A	25	p.Gln742Lys	Castella, 2011b	1	1
c.2426G>A	26	p.Gly809Asp	Rockfeller's FA Database	1	1
c.2574C>G	27	p.Ser858Arg	Tamary, 2000	1	1
c.2639G>A	28	p.Arg880Gln	Rockfeller's FA Database	1	1
c.2851C>T	29	p.Arg951Trp	Chandra, 2005	1	1
c.3164G>T	32	p.Arg1055Leu	Adachi, 2002	1	1
c.3239G>A	32	p.Arg1080Gln	Castella, 2011b	1	1
c.3335T>G	33	p.Val1112Gly	Castella, 2011b	1	1
Inserciones/delecciones: 39,1%					
Cambio en DNA	Exón	Proteína	Referencia	Alelos	Pacientes
c.3788_3790delTCT	38	p.Phe1263del	Magdalena, 2005	29	23
c.1115_1118delTTGG	13	p.Val372Alafs*42	FA/Breast cancer consortium, 1996	24	20
c.3558dupG	36	p.Arg1187Glufs*28	Savino, 1997	17	15
c.233_236delTTGA	3	p.Thr78Ilefs*16	Castella, 2011b	3	3
c.238delT	3	p.Cys80Valfs*15	Castella, 2011b	2	1
c.472insC	5	p.His158Profs*3	No reportada	2	2
c.1164_1165delAG	13	p.Arg388Serfs*20	Wijker, 1999	2	2
c.1703delT	18	p.Val568Alafs*37	Castella, 2011b	2	2
c.2535_2536delCT	27	p.Cys846Glnfs*20	Ameziane, 2008	2	2
c.50delG	1	p.Gly17Alafs*27	Castella, 2011b	1	1
c.284delG	4	p.Gly95Alafs*43	Castella, 2011b	1	1
c.987_990delTCAC	11	p.His330Alafs*4	Morgan, 1999	1	1
c.1858_1859insC	21	p.Tir621Leufs*9	No reportada	1	1
c.2233dupT	25	p.W745Lfs*49	Rockfeller's FA Database	1	1
c.3446insC	35	p.Leu1150Profs*63	No reportada	1	1
c.4124_4125delCA	41	p.Thr1375Serfs*49	Rockfeller's FA Database	1	1
c.4123_4139del17	41	p.Thr1375Trpfs*44	Castella, 2011b	1	1
c.4253delT	42	p.V1418Gfs*15	No reportada	1	1
Splicing: 7,2%					
Cambio en DNA	Intrón	Proteína	Referencia	Alelos	Pacientes
c.3828+1G>T	38i	ND	Castella, 2011b	4	4
c.1083+2T>C	12i	ND	No reportada	2	2
c.3348+5G>A	33i	ND	Castella, 2011b	2	2
c.893+1G>T	10i	ND	Savino, 1997	2	2
c.1470+1G>T	15i	ND	Castella, 2011b	1	1
c.3348+18A>G	33i	p.[(Glu1106Lys; Asn1113Ile)]	Rockfeller's FA Database	1	1
c.709+5G>A	7i	p.Asp237Gly+ins10 (AFMTRCGFLD)	Ameziane, 2008	1	1
c.709+5G>T	7i	p.(Asp237_Phe238delinsGAFMTRCGFLD)	Lo Ten Foe, 1997	1	1
c.710_5T>C	7i	ND	Rockfeller's FA Database	1	1
c.893+2T>C	10i	ND	No reportada	1	1
c.893+920C>A	10i	ND	Castella, 2011b	1	1
Delecciones intragénicas: 11,9%					
Delección	Alelos	Pacientes			
Exon 3	4	2			
Exones 1-20	2	2			
Exones 1-3	1	1			
Exones 15-17	1	1			
Exones 15-43	1	1			
Exones 1-6	1	1			
Exones 16-17	1	1			
Exones 18-20	1	1			
Exones 19-21	1	1			
Exones 19-22	1	1			
Exones 21-22	1	1			
Exones 21-29	3	3			
Exones 27-36	1	1			
Exones 28-29	1	1			
Exones 29-30	1	1			
Exones 32-33	1	1			
Exones 33-36	1	1			
Exones 4-5	2	2			
Exones 4-6	2	1			
Exones 6-14	1	1			

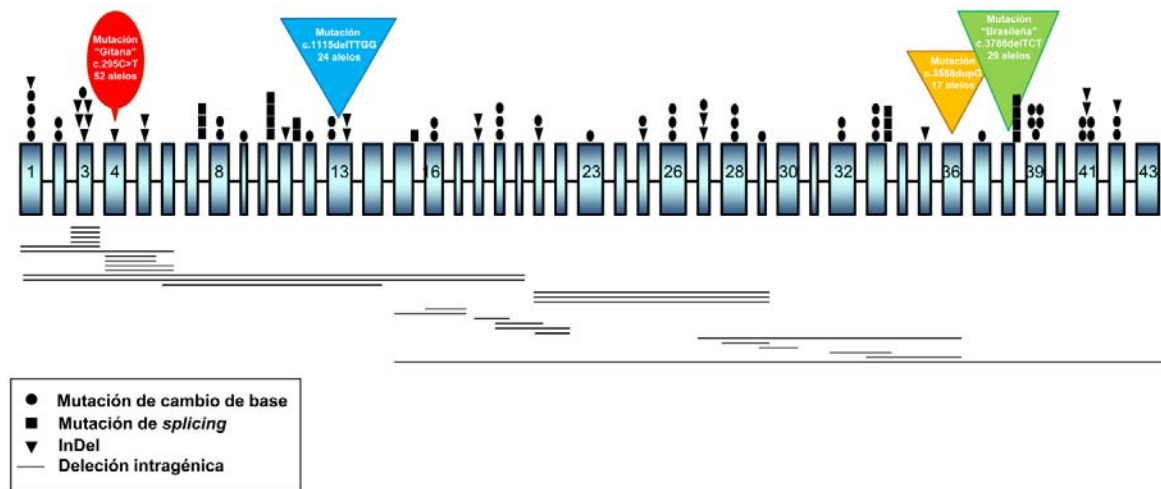


Figura 29: Esquema del gen *FANCA* y de las mutaciones encontradas en pacientes de la población española. Los bloques representan los exones y las líneas que los unen los intrones. Los alelos marcados en rojo corresponden a los %2 alelos identificados de la mutación c.295C>T característica de la etnia gitana. Todos los otros alelos se han encontrado en pacientes de etnia caucásica. En azul se muestra la mutación c.1115delTTGG presente con 24 alelos, en amarillo la c.3558dupG presente con 17 alelos y en verde la mutación “Brasileña” c.3788delTCT (Magdalena et al., 2005) presente con 29 alelos (Modificada de (Castella et al., 2011b)).

VI.1.3. Espectro mutacional de los pacientes no FA-A analizados

Las mutaciones encontradas en los pacientes pertenecientes a otros grupos de complementación muestran que existe también en estos genes un espectro mutacional amplio dentro de la cohorte analizada (Tabla XII). Se detectó una elevada proporción de mutaciones nuevas (8/25). Las mutaciones bialélicas detectadas en *BRCA2/FANCD1* (2 pacientes) y en *BRIPI/FANCI* (1 paciente) son de particular relevancia ya que confieren a los portadores un incremento del riesgo de desarrollar un cáncer de mama/ovario. Los métodos de caracterización molecular tradicionales de los pacientes FA consideran que los genes *FANCA*, *C* y *G* son los más frecuentemente mutados en la población general FA basándose en los datos de los pacientes de Estados Unidos (EEUU) (Auerbach, 2009). Sin embargo esto no es correcto para la población española donde el segundo gen más frecuentemente mutado es *FANCD2* (7%) (Figura 27) mientras que las mutaciones en *FANCC* no tienen casi relevancia (0,64%, un solo paciente identificado en España). La casi totalidad de los pacientes FA de EEUU con mutaciones en el gen *FANCC* pertenece a la etnia de los judíos Ashkenazi que, por efecto fundador, tienen una prevalencia de portadores de 1/100 de la mutación c.711+4A>T en *FANCC* (Kutler and Auerbach, 2004).

Tabla XII: Espectro mutacional de los pacientes NO FA-A

FANCB				
Cambio en DNA	Exón/intrón	Cambio en Proteína	Referencia	Nº de pacientes
c.786G>T	3	p.Lys262Asn	No reportada	1
FANCD1/BRCA2				
Cambio en DNA	Exón/intrón	Cambio en Proteína	Referencia	Nº de pacientes
c.7796A>G	16	p.Glu2599Gly	No reportada	1
c.1814insA	10	Pro606Thrfs*8	rs80359310	1
c.8488-1G>A	i19		Howlett et al., 2002	1
FANCD2				
Cambio en DNA	Exón/intrón	Cambio en Proteína	Referencia	Nº de pacientes
c.98InsA;	3	p.Lys34Glufs*6	No reportada	1
c.2444G>A	26	p.Arg815Gln;	Kalb et al., 2007	2
c.3777+1G>A	i37		No reportada	1
c.1068T>A	13	p.Tyr356*	No reportada	1
c.226_228delA	3	p.Lys77Argfs*7	No reportada	2
FANCE				
Cambio en DNA	Exón/intrón	Cambio en Proteína	Referencia	Nº de pacientes
c.929insC	4	p.Val311Serfs*1	No reportada	1
FANCG				
Cambio en DNA	Exón/intrón	Cambio en Proteína	Referencia	Nº de pacientes
c.1642C>T	13	p.Arg548*	Ahuerbach, 2003	2
c.450G>A		p.Trp150*		1
c.566dupA	5	p.Glu190Glyfs*45	No reportada	1
FANCI				
Cambio en DNA	Exón/intrón	Cambio en Proteína	Referencia	Nº de pacientes
c.751C>T	7	p.Arg251Cys	Chandrasekharappa,2013	1
c.975_977delA	7	p.Arg279Glyfs*10	No reportada	1
FANQ/ERCC4				
Cambio en DNA	Exón/intrón	Cambio en Proteína	Referencia	Nº de pacientes
c.1484_1488delCTCAA	8	p.Thr495Asnfs*5	Bogliolo et al., 2013	1
c.2065C>A	11	p.Arg689Ser	Bogliolo et al., 2013	1

Por lo tanto, la detección de mutaciones en el gen *FANCD2* debería optimizarse para la correcta caracterización molecular de los pacientes FA Españoles de etnia no gitana. La presencia de dos pseudogenes complica el desarrollo de técnicas sensibles para detectar las deleciones intragénicas y las variantes intrónicas a nivel del DNA genómico de *FANCD2* (Kalb et al., 2007). Actualmente estamos desarrollando métodos basados en el análisis del mRNA de *FANCD2* para caracterizar los casos más complicados como el caso de los pacientes FA664 y FA681. FA664 y FA681 son hermanos y ambos FA mosaicos y en el DNA de PB logramos identificar por WES solo uno de los alelos mutados de *FANCD2*. En el caso del paciente FA681 al haber disponible una línea de fibroblastos primarios, pudimos

estudiar FANCD2 por WB: los fibroblastos del paciente no presentaban banda visible correspondiente a FANCD2 confirmando así el dato de la WES que nos hacía sospechar que el subtipo fuese FA-D2. No obstante se volvió a repetir el análisis por WES en DNA de fibroblastos primarios del paciente FA681 y no se consiguió identificar la segunda mutación en *FANCD2*, lo que hace pensar que la mutación esté muy adentro de una de las secuencias intrónicas o que sea una gran deleción.

VI.1.3.1 El caso del paciente FA663

El paciente FA663 es una mujer que nació con bajo peso, eminencia tenar plana y hiperpigmentación de la piel. Los valores hematológicos eran normales al momento del nacimiento pero a lo largo de la vida se le detectó una macrocitosis de origen desconocido. No había ninguna información sobre de los datos biológico de los padres porque la paciente había sido adoptada. A los 33 años se le diagnosticó un cáncer de mama ductal invasivo. Se le practicó la mastectomía de la mama derecha y después del primer ciclo de quimioterapia la paciente manifestó una pancitopenia muy severa. Con la sospecha clínica de FA fue derivada a nuestro servicio de diagnóstico y en el test de fragilidad cromosómica la paciente resultó tener una fragilidad cromosómica compatible con un diagnóstico de FA mosaico (32% de células aberrantes). Su caso fue incluido en nuestra cohorte para ser caracterizado molecularmente con el método WES + MLPA y se descubrieron dos variantes en el gen *BRCA2/FANCD1*: la primera variante (c.1813_1814insA, p.Pro606Thrfs*10) era una mutación en el exón 10 que producía un cambio de pauta de lectura y la terminación prematura de la traducción y, por tanto, era claramente patogénica. Además esta mutación había sido previamente descrita en pacientes FA-D1 (Wagner et al., 2004) y en pacientes con cáncer de mama familiar (*Breast Cancer Information Core (BIC)*). La segunda era una variante de cambio de sentido (c.7796A>G, p.Glu2599Gly) no reportada previamente pero potencialmente perjudicial según el análisis *in silico* (Tablas VII). Para confirmar que ambas las variantes encontradas en *BRCA2* eran patogénicas, se le practicó una biopsia de piel y se estudió la formación de *foci* de RAD51 después de tratamiento con MMC en los fibroblastos primarios. Se vio que la formación de *foci* de RAD51 después de tratamiento con MMC era defectiva y esta observación concordaba con un defecto de la HRR que sugería que ambos

alelos de *BRCA2/FANCD1* no eran funcionales. En base a nuestro diagnóstico se recomendó una mastectomía profiláctica también de la mama izquierda. A pesar de que antes de la mastectomía el pecho izquierdo parecía normal según los análisis ecográficos y radiológicos, en el sucesivo análisis histopatológico se detectó una forma muy agresiva de carcinoma ductal. Afortunadamente el carcinoma todavía era pequeño e *in situ*.

El caso de la Paciente F663 es de peculiar importancia por dos razones: la primera es que la rapidez y la fiabilidad del método WES+MLPA con toda probabilidad salvaron la vida de la paciente y la segunda razón es que la paciente FA663 es el único paciente FA conocido hasta el momento con mutaciones bialélicas en *BRCA2/FANCD1* que tenga un fenotipo tan suave (Alter et al., 2007, Wagner et al., 2004, Meyer et al., 2014). Esto es probablemente debido a que la mutación c.7796A>G, p.Glu2599Gly tiene un marcado carácter hypomórfico y porque la reversión de una de las mutaciones en el tejido hematopoyético tuvo lugar antes que la paciente desarrollara un fenotipo FA grave a nivel hematológico.

VI.1.3.2 El caso del paciente FA574

El paciente FA574 es un varón que nació en 1992 y en el 2007 empezó a mostrar síntomas de BMF. El test de fragilidad en PB dio positivo y el diagnóstico fue de FA mosaico (21% de células aberrantes después de tratamiento con DEB). La complementación retroviral descartó mutaciones en los genes *FANCA*, *FANCC*, y *FANCG*. El análisis por WB de los fibroblastos primarios del paciente reveló que no había monoubiquitinación de FANCD2 después de tratamiento con MMC y esto era compatible con una alteración de la función de un gen perteneciente al FANCore. El análisis por WES+MLPA detectó una variante de cambio de sentido (c.786G>T; p.Lys262Asn) en el exón 3 de *FANCB* que en el análisis *in silico* resultó ser probablemente patogénica (Tabla VII). El grupo de complementación FA-B es muy raro y *FANCB* tiene la peculiaridad de ser el único gen FA conocido hasta al momento por localizarse en el cromosoma X (Meetei et al., 2004, Holden et al., 2006, McCauley et al., 2011). Por tanto analizamos por PCR y secuenciación de Sanger el exón 3 de *FANCB* en todos los otros miembros de la familia: la madre, el padre, un hermano sano y una hermana sana. La variante se detectó en heterocigosis en la madre pero

no se detectó en el padre y en los dos hermanos sanos. La segregación por lo tanto es compatible con la hipótesis que el paciente FA574 sea el primer paciente español FA-B.

VI.2. Identificación de *ERCC4* como gen FA

En el paciente FA104, que había sido extensivamente estudiado sin encontrar el gen mutado (Casado et al., 2007), identificamos por WES mutaciones bialélicas en *ERCC4*. La identificación de otro paciente FA no emparentado (el paciente alemán 1333) con mutaciones bialélicas en *ERCC4* definitivamente confirmó este gen como un nuevo gen FA y lo rebautizamos con el alias de *FANCO* con la aprobación del *HUGO Gene nomenclature committee* (Bogliolo et al., 2013). *ERCC4* ya se había identificado en 1996 como el gen defectuoso en los pacientes de XP pertenecientes al grupo F y desde entonces la proteína se conoce como XPF (Lehmann et al., 1994, Sijbers et al., 1996). Los genes *ERCC4* y *ERCC1* codifican proteínas que se asocian en un dímero con actividad endonucleasa que participa en la reparación del DNA y en el mantenimiento de la estabilidad genómica. El complejo ERCC1-XPF juega un papel central en la NER, un proceso que elimina los daños en el DNA causados por UV (CPDs y 6-4PPs) y por agentes que producen aductos que distorsionan la doble hélice del DNA (Scharer, 2013) (Figura 30). La formación del complejo ERCC1-XPF es también necesaria para la estabilidad de las dos proteínas: cuando ERCC1 o XPF se expresan en ausencia de la otra forman agregados que se degradan por proteólisis (Gaillard and Wood, 2001). ERCC1-XPF, gracias a la actividad catalítica de XPF, corta el DNA dañado en el lado 5' de la "burbuja" de DNA abierto formada durante la NER. Después del corte en 3' por parte de la otra endonucleasa XPG, se produce la escisión de un fragmento de ssDNA de 25-30 nucleótidos que incluye las bases dañadas (Sijbers et al., 1996, Evans et al., 1997, Mu et al., 1996, Scharer, 2013) (Figura 27). Defectos hereditarios en la NER son causa de XP, que se caracteriza por fotosensibilidad extrema y por un aumento de 1000 veces del riesgo de padecer cáncer de piel (Friedberg et al., 2006b). Defectos genéticos en algunas de las proteínas de la NER producen también otros trastornos: debido a las funciones adicionales que XPB y XPD tienen en la transcripción como subunidades del factor II de la transcripción (TFIIH) defectos en estas proteínas además de XP pueden causar síndrome de Cockayne (CS), síndrome combinado XP/Cockayne (XP/CS) y Tricotodistrofia (TTD),

caracterizados por defectos graves del desarrollo y neurodegeneración (Lehmann, 2003). Mutaciones que afectan a XPG, la otra endonucleasa involucrada en la NER, pueden también producir XP/CS (Nouspikel et al., 1997, Cleaver et al., 1999). Mutaciones en ERCC1 causan el síndrome cerebro-óculo-facio-esquelético (COFS) (Jaspers et al., 2007) y CS (Kashiyama et al., 2013), mientras que mutaciones en XPF además de XP pueden producir el síndrome de envejecimiento acelerado XFE (Niedernhofer et al., 2006), CS (Kashiyama et al., 2013) y un fenotipo de XP/CS asociado con pancitopenia (Kashiyama et al., 2013).

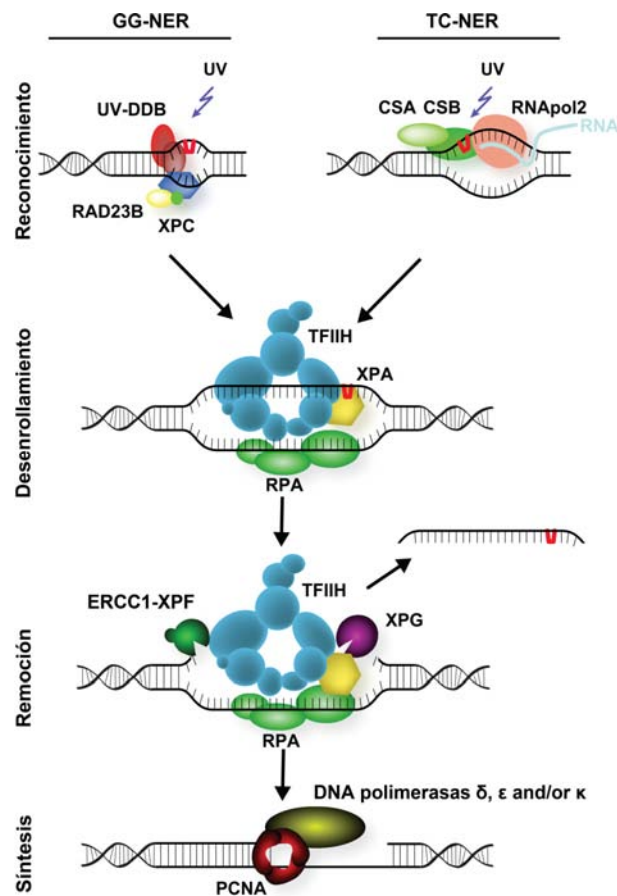


Figura 27: NER en mamíferos: La NER es activada por dos mecanismos diferentes de detección de daños: la NER acoplada a la transcripción (TC-NER) se inicia cuando la RNA polimerasa II se bloquea en proximidad de una lesión presente en la hebra transcrita de los genes y depende de reclutamiento de las proteínas CSA y CSB mutadas en CS. Las lesiones en el DNA que no se transcribe de forma activa se eliminan por NER del genoma global (GG-NER), que se inicia por la ubiquitina ligasa UV-DDB de y el complejo XPC-RAD23-CETN2. Después de la detección, se recluta a la lesión el factor de transcripción TFIIH que desenrolla un tramo de aproximadamente 30 nucleótidos proporcionando el acceso a otros factores de la reparación como las proteínas XPA y RPA que estimulan la actividad de TFIIH y de las endonucleasas ERCC1-XPF y XPG para que corten el DNA. Después de la escisión de la cadena dañada la doble cadena se restablece gracias a la fase de síntesis y de ligación del DNA. Esta fase involucra factores de la replicación como PCNA y RFC, las DNA polimerasas δ, ε y κ, y las DNA ligasas I y III. Para mayor claridad, no se muestran todas las proteínas que se sabe estar involucradas en la NER (Modificada de (Lans et al., 2012)).

Los ratones *Ercc1*^{-/-} nacen con prevalencias cerca de las mendelianas esperadas pero su crecimiento postnatal es extremadamente retrasado y los ratones mueren aproximadamente a las 3 semanas de edad cuando pesan solo un 20% en comparación con ratones normales (McWhir et al., 1993, Weeda et al., 1997). Además desarrollan espontáneamente síntomas característicos de neurodegeneración progresiva (Niedernhofer et al., 2006) y, aunque el sistema hematopoyético se desarrolle normalmente, al final de la vida los ratones *Ercc1*^{-/-} tienen leucopenia y trombocitopenia (Prasher et al., 2005).

Los ratones *Xpf*^{-/-}, como en el caso de los ratones *Ercc1*^{-/-}, se desarrollan normalmente pero el crecimiento postnatal se retrasa de tal manera que a las 2 semanas los ratones *Xpf*^{-/-} son aproximadamente el 25% del tamaño normal y mueren a las 3 semanas de edad (Tian et al., 2004). Los paralelos fenotípicos entre los ratones *Ercc1*^{-/-} y *Xpf*^{-/-} sugieren que las dos proteínas funcionan exclusivamente como un complejo.

Las células nulas para XPF y ERCC1 son excepcionalmente sensibles a los agentes que inducen ICL como el cisplatino o la MMC (Wood, 2010) mientras que las células defectuosas en otros genes involucrados en la NER son comparativamente menos sensibles (Andersson et al., 1996, Damia et al., 1996). Esta característica específica de ERCC1-XPF hizo pensar desde el primer momento que esta endonucleasa pudiese estar involucrada en la reparación de lo ICL y que este papel podía ser independiente de la NER (Ahmad et al., 2008, Fisher et al., 2008, Wood, 2010, Rahn et al., 2010, Kuraoka et al., 2000, Clingen et al., 2005, Al-Minawi et al., 2009). Las funciones en la NER y en la ICLR de ERCC1-XPF se pueden separar artificialmente mutando aminoácidos de ERCC1 importantes para la interacción con XPA: las células que expresan estas formas mutadas de ERCC1 son sensibles a la radiación UV pero no a los ICL (Orelli et al., 2010).

Nuestra hipótesis para explicar porqué los pacientes FA104 y 1333 tenían fenotipo clínico FA y no XP era que las mutaciones encontradas en XPF, de forma parecida a las mutaciones introducidas artificialmente en ERCC1, separaban de forma natural las funciones de ERCC1-XPF en la NER y en la ICLR. En esta tesis hemos conseguido demostrar de forma indiscutible que las mutaciones en XPF que producían FA inhibían por completo la capacidad de ERCC1-XPF de actuar en la ICLR pero solo afectaban marginalmente su función en la NER. Por lo tanto esto explicaba porqué los pacientes FA104 y 1333 no tenían

rasgos de XP u otras enfermedades relacionadas con defectos de la NER (Figuras 17-21). Cuando las mutaciones en XPF perjudican a ambas vías de reparación, se producen enfermedades más complejas como el síndrome progeroide XFE. El único paciente de este síndrome presentaba una plétora de manifestaciones clínicas entre las cuales había extrema fotosensibilidad y anemia (Niedernhofer et al., 2006) (Figura 28). Los fibroblastos primarios derivados del paciente XFE tienen una extrema sensibilidad a la radiación UV (Ahmad et al., 2010) y demuestran una fragilidad cromosómica típica de FA en respuesta a los ICL (Figura 21). Por contrario los fibroblastos primarios de los pacientes XPF no tienen fragilidad cromosómica en respuesta a los ICL pero son extremadamente sensibles a los UV (Figura 23). Por lo tanto, nuestra hipótesis quedaba confirmada por completo: mutaciones en XPF que inactivan la NER producen XP, mutaciones que inhiben la ICLR producen FA y mutaciones que inhiben las dos rutas de reparación pueden producir XFE, un síndrome que combina muchas de las características clínicas y moleculares de FA y XP-F (Tabla XIII y Figura 28). Los pacientes XFE tienen rasgos de fotosensibilidad típicos de XP-F y retraso del desarrollo y anemia típicos de FA. Hay pero que remarcar que no se sabe si la anemia del paciente XFE era causada por BMF. A nivel celular la combinación de estas características se hace aún más evidente. Las células XFE son sensibles a UV tanto a nivel de la sobrevivencia celular como a nivel de la UDS y también tienen una sensibilidad a los ICL idéntica a la de las células FA (Figura 23 y Tabla XIII).

Tabla XIII: Características clínicas y fenotipos celulares de los distintos síndrome producidos por mutaciones bialélicas en ERCC4

Características clínicas y celulares	XP-F	XFE	FA-Q
Sensibilidad de la piel a la luz	Leve	Severa	No
Epidermis atrófica	Variable	Si	No
Rasgos neurológicos	Poco frecuentes	Si	No
Hematología	Normal	Anemia ^a	Anemia, BMF
Retraso del desarrollo ^b	No	Si	Si
Fallecimiento prematuro	No	16 años	4 años (FA104). 1333 vivo a los 10 años
Sensibilidad a los UV	Leve	Severa ^c	Ninguna (FA104) ^d , Leve(1333)
Disminución de la UDS	Leve	Severa ^c	Leve (1333), ND en (FA104) ^d
MMC sensitivity	Leve	Severa	Severa
DEB-test	Negativo	Positivo	Positivo
MMC induced G2/M arrest	Negativo	Positivo	Positivo
Actividad nucleasica <i>in vitro</i>	Si ^e	Si ^e	No (FA104)

La características típicas de XP-F se evidencian en color amarillo y las de FA-Q en verde. ^aSe desconoce si hubo BMF en el paciente XPF (Laura Niedernhofer, comunicación personal). ^bMicrosomía en los pacientes 1333, FA104 y XFE y microcefalia en los pacientes XFE y 1333. ^c(Niedernhofer et al., 2006). ^dLa UDS no se testó en el paciente FA104 por la falta de fibroblastos primarios de piel pero los linfoblastos eran resistentes a los UV. ^e(Ahmad et al., 2010).

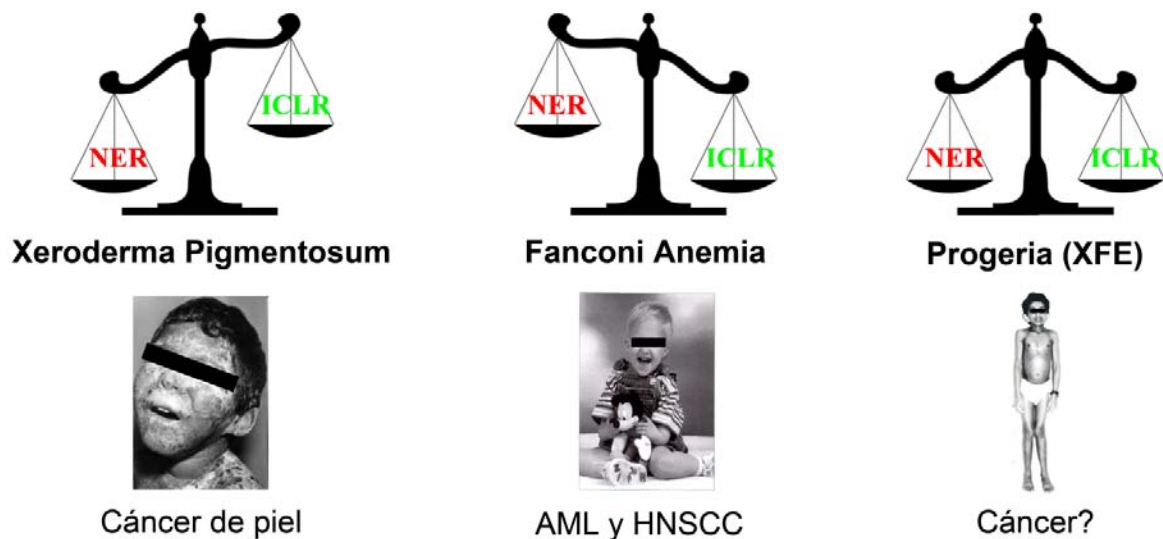


Figura 28: Equilibrio entre la NER e ICLR en las enfermedades humanas causadas por mutaciones en XPF: mutaciones que inhiben la NER producen XP, mutaciones que inhiben la ICLR producen FA y mutaciones que afectan a ambas vías producen XFE.

Tratamos también de investigar a través de que mecanismo molecular se producía esto desacoplamiento de la funciones de ERCC1-XPF y estudiamos su interacción específica de la ICLR con SLX4 (Crossan et al., 2011, Cybulski and Howlett, 2011, Kim et al., 2011, Klein Douwel et al., 2014, Stoepker et al., 2011). Las formas mutadas de XPF que producían FA eran capaces de entrar normalmente en el núcleo, de interactuar con SLX4 y de relocalizar a la cromatina (Figura 24 A y D). Por lo tanto el defecto en la ICLR podía estar en la fase de desenganche debido a un defecto de XPF en la capacidad de cortar el DNA. La proteína XPF-p.Arg689Ser se demostró activa, como se esperaba, en el ensayo de NER *in vitro* (Figura 25 A) pero los productos de escisión más abundantes eran de media un nucleótido más largos respecto a los productos de escisión de XPF-wt. En el ensayo de corte de una estructura tallo-bucle (Figura 25 B) la XPF-p.Arg689Ser era totalmente incapaz de cortar el DNA. Esto demostraba que, a pesar de que el mutante XPF-p.Arg689Ser retuviera suficiente capacidad de escindir el DNA como para conseguir una reacción de NER casi normal, algo fallaba en su capacidad de cortar determinadas estructuras del DNA. Lamentablemente solo pudimos realizar los estudios de la actividad catalítica del mutante XPF-p.Arg689Ser porque no fue posible purificar la proteína XPF-p.Leu230Pro por su baja

estabilidad. Decidimos por lo tanto estudiar con la técnica del COMET (De Silva et al., 2000, De Silva et al., 2002) el desenganche de los ICL en MEF deficientes para Xpf que expresaban las formas mutadas XPF-p.Arg230Ser y XPF-p.Leu230Pro con (Figura 25 C). Sorprendentemente las células que expresaban las dos formas mutantes eran capaces de cortar el DNA dañado por ICL a niveles muy parecidos a los de las células que expresaban el XPF-wt. Una posible explicación de nuestras observaciones era que, como en el caso del ensayo NER *in vitro*, el corte del DNA se produjera pero no de la forma correcta dejando así un sustrato no procesable por las etapas posteriores a la fase de desenganche. Esto concuerda con lo observado en células FA-A que tienen capacidad de desenganche normal (Rothfuss and Grompe, 2004) pero estas observaciones discrepan con lo observado en sistemas de reparación *in vitro* de los ICL donde la falta de FANCD2 o de FANCI o de su monoubiquitinación inhibe la fase de desenganche de la ICLR (Knipscheer et al., 2009). Por lo tanto aún no está del todo claro qué mecanismo molecular explica porqué las mutaciones p.Arg689Ser y p.Leu230P en ERCC4 solo producen FA. Mutaciones que se localizan entre los aminoácidos 225 y 236 del dominio helicasa de XPF pueden producir XP, FA, CS y XP/CS/FA (Figura 29) y, por lo tanto, el dominio helicasa, aunque no funcional (Ciccia et al., 2008), es fundamental para la función y la regulación de XPF en la reparación del DNA y en enfermedades humanas. Es posible que esta regulación tan fina sea el producto de interacciones con otras proteínas específicas de las distintas vías de reparación del DNA y que mutaciones específicas en XPF puedan afectar algunas interacciones dejando a otras en condiciones relativamente normales. Variaciones en el interactoma de XPF en función de mutaciones específicas podrían así explicar porqué mutaciones a veces en el mismo dominio de XPF (Figura 29 y Figura 30) puedan causar hasta cinco enfermedades (XP, FA, CS, Progeria XFE y XP/CS) fenotípicamente muy diferentes y que tienen también características de predisposición al cáncer distintas. De hecho los pacientes FA desarrollan AML y HNSCC (Bogliolo and Surralles, 2015), los pacientes de XP son muy propensos al cáncer de piel y los pacientes CS no muestran predisposición al cáncer (Lehmann, 2003). Estudios proteómicos y funcionales en curso en nuestro grupo podrán quizá desvelar el interactoma de las diferentes formas mutantes de XPF y ayudar a desentrañar nuevos mecanismos implicados en el mantenimiento de la estabilidad del genoma, descubrir nuevas respuestas celulares a los medicamentos contra el cáncer o identificar nuevos biomarcadores y dianas terapéuticas.

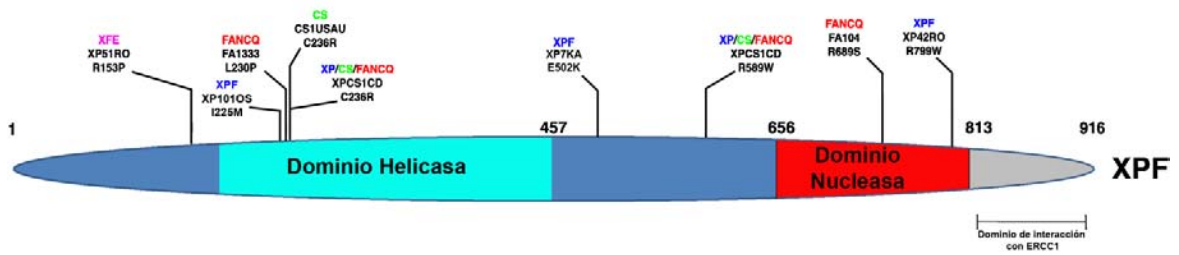


Figura 29: Esquema de la proteína XPF con ejemplos de mutaciones encontradas en diferentes síndromes y las líneas celulares derivadas de los pacientes (Modificada de (Manandhar et al., 2015)).

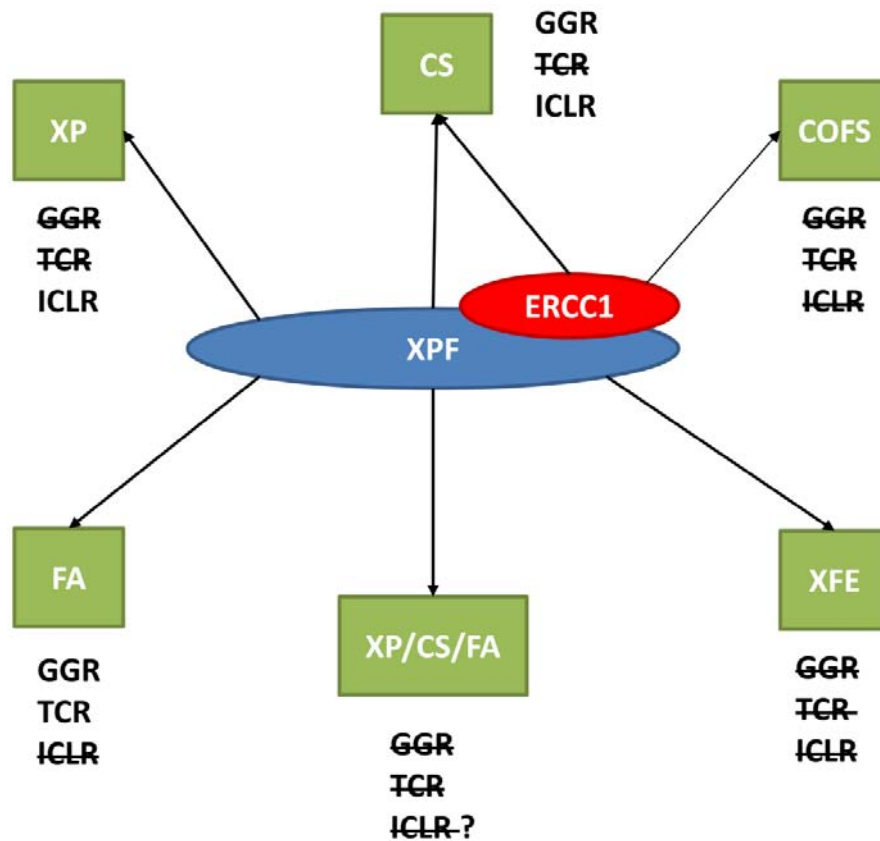


Figura 30: Mutaciones que afectan el papel de la endonucleasa XPF-ERCC1 en las distintas vía de reparación del DNA producen hasta seis enfermedades diferentes.

VI.3. *ERCC4* y FBOC

Como ya se ha dicho en el apartado “Introducción” de esta tesis, tradicionalmente solo los genes FA directamente involucrados en la HRR se solían asociar con la predisposición a FBOC. Defectos en la HRR (sensibilidad a la Camptotecina y a los inhibidores de la PARP) asociados a mutaciones bialélicas de un gen FA se solían considerar como marcadores que dicho gen podía ser también un candidato para la susceptibilidad a FBOC. Cuando se descubrió que mutaciones bialélicas en *SLX4* producían FA, todo hacía presagiar que *SLX4* sería un gen importante para la predisposición a FBOC al ser *SLX4* importante en la HRR y siendo las células deficientes para *SLX4* sensibles a la Camptotecina y a los inhibidores de la PARP (Fekairi et al., 2009, Munoz et al., 2009, Saito et al., 2009, Svendsen et al., 2009). Pero no fue así (Landwehr et al., 2011, Catucci et al., 2012, Fernandez-Rodriguez et al., 2012, Bakker et al., 2013, de Garibay et al., 2013, Shah et al., 2013). Además las células FA-J con mutaciones bialélicas en *BRIP1*, gen importante para la susceptibilidad a FBOC, no muestran defecto alguno en la HRR y por lo tanto al no haber ningún ensayo funcional que nos permita hacer una criba, cualquier nuevo gen FA es un excelente candidato para explicar la susceptibilidad a FBOC de familias negativas para mutaciones en los genes conocidos hasta al momento. Para investigar el posible papel de *ERCC4* como gen de susceptibilidad a FBOC estudiamos en colaboración con el Grupo de Genética Humana del Centro Nacional de Investigaciones Oncológicas (CNIO) de Madrid los 11 exones y las regiones intrónicas proximales del gen *ERCC4* en 1573 casos españoles de alto riesgo de FBOC y negativos para mutaciones en *BRCA1/FANCS* y *BRCA2* y en 854 individuos controles. La prevalencia de portadores de mutaciones *ERCC4* no difiere entre los casos y los controles ($p = 0,251$), lo que sugiere que las mutaciones monoalélicas en *ERCC4* no están vinculadas a la susceptibilidad al cáncer en la población en general.

Se identificaron cinco variantes únicas entre los casos y cuatro en los controles y siete de estas se consideraron como supuestamente perjudiciales. Las mutaciones sin sentido se evaluaron utilizando la herramienta bioinformática CONDEL (Gonzalez-Perez and Lopez-Bigas, 2011) y se estudiaron funcionalmente las variantes que se predijeron afectar a la función de XPF. En cuanto a los casos, una de las mutaciones encontradas, c.584+1G> A en el intrón 3, se confirmó que producía la pérdida del exón 3 en el mRNA causando un cambio de la pauta de lectura y una terminación prematura de la traducción (no mostrado) y las otras

dos mutaciones la p.Arg150Cys en el exón 3, y la p.Ser786Phe en el exón 11, se investigaron funcionalmente. En cuanto a los controles, se identificaron dos mutaciones la c.540_541delAG en el exón 3 (que se encuentra en dos controles independientes) y la c.2291delG en el exón 11 que producen cambio del marco de lectura y un fin anticipado de la traducción. La última variante deletérea encontrada en los controles fue el cambio de sentido p.Arg689Ser en el exón 11 que se había encontrado previamente en el paciente de FA FA104 (Bogliolo et al., 2013).

En conclusión, la prevalencia de individuos heterocigóticos españoles para las mutaciones deletéreas en *ERCC4* es de aproximadamente el 0,3% y sugiere que en la población española se deberían encontrar según Hardy–Weinberg, un centenar de individuos con mutaciones bialélicas dañinas en *ERCC4* pero, de momento, solo se identificó a un paciente (Bogliolo et al., 2013). Las prevalencias de cada uno de los grupos de complementación de FA se considera que deriven de las prevalencias de portadores en la población general, pero se desconoce si mutaciones bialélicas en genes FA estén asociadas a mayor letalidad embrionaria. Esto pero no es cierto en el caso del gen *BRCA2/FANCD1*: se calcula que la prevalencia de portadores de mutaciones monoalélicas en la población general es de 1/500 (Schwartz et al., 2008), por lo que en España esperaríamos de encontrar alrededor de 40 casos FA-D1, pero hasta el momento solo se descubrieron 4. Por lo tanto se puede afirmar que, de forma similar a *BRCA2/FANCD1*, la mayoría de las combinaciones de mutaciones bialélicas en *ERCC4* producen letalidad embrionaria. Además todos los pacientes XP-F reportados tienen al menos una mutación sin sentido (Cleaver et al., 1999) mientras que todas las mutaciones sin sentido que se encuentran en los 2.427 individuos españoles tienen una actividad NER normal, lo que explica por qué no haya sido identificada hasta al momento ninguna familia XP-F en España.

CONCLUSIONES

“Everything will be okay in the end. If it’s not okay, it’s not the end”.

Anónimo

VII. COCLUSIONES

1. El método WES+MLPA para la caracterización de los pacientes FA se ha revelado funcionar de forma excelente y en este momento se puede considerar el mejor sistema en términos de rapidez, sensibilidad y costes tanto para la determinación del subtipo como de las mutaciones en los pacientes FA.
2. Los kits de captura comerciales para el exoma entero tienen suficiente cobertura de los genes FA para permitir la detección fiable de las mutaciones en todos los genes conocidos hasta al momento.
3. Los datos de cobertura generados por WES se pueden utilizar para detectar deleciones intragénicas. En un futuro próximo la MLPA podría ya no ser necesaria para la caracterización molecular de los pacientes FA.
4. Para la caracterización de los pacientes mosaicos con el método WES+MLPA es aconsejable la utilización de DNA de fibroblastos primarios cultivados a partir de biopsias de piel.
5. En la población FA española el segundo grupo de complementación por número de pacientes es el FA-D2. Esto es diferente respecto a la población de pacientes FA de EEUU y se necesitaría la puesta en marcha de sistemas complementarios a la WES para la detección de las mutaciones en el gen FANCD2 como, por ejemplo, la secuenciación del cDNA.
6. Mutaciones bialélicas en *ERCC4* causan FA y *ERCC4* es por tanto el responsable de un nuevo grupo de complementación FA-Q y se rebautiza con el alias de *FANCO*.
7. Las mutaciones en *ERCC4* que inhiben la actividad de XPF en la ruta de la ICLR y no afectan a su papel en la NER, causan FA.

8. Mutaciones que afectan solo a la actividad de XPF en la NER causan XP mientras que mutaciones que inhiban tanto la NER como la ICLR producen síndromes más complejos como XFE.
9. Las proteínas XPF alteradas en FA tienen una actividad nucleasa casi normal en la NER pero no pueden cortar determinados sustratos de DNA.
10. Las proteínas XPF mutadas en FA son capaces de interactuar con SLX4, llegar al sitio del daño en el DNA y de efectuar el “unhooking” de los ICL. Los problemas que presentan a la hora de cortar determinado sustratos in vitro hace pensar que el desenganche del ICL no se efectúe de la forma correcta y deje así un sustrato que no se puede ulteriormente procesar.
11. Mutaciones monoalélicas en *ERCC4* son relativamente frecuentes en la población española, pero no causan predisposición genética al cáncer de mama y ovario.
12. La identificación de un solo paciente con mutaciones bialélicas en *ERCC4* en la población española donde hay una prevalencia de portadores del 0,3% respalda la hipótesis que la mayoría de las combinaciones de mutaciones bialélicas en *ERCC4* sean letales a nivel embrionario.

BIBLIOGRAFÍA

“Think before you speak. Read before you think.”

Fran Lebowitz

VIII. BIBLIOGRAFÍA:

- ADACHI, D., ODA, T., YAGASAKI, H., NAKASATO, K., TANIGUCHI, T., D'ANDREA, A. D., ASANO, S. & YAMASHITA, T. 2002. Heterogeneous activation of the Fanconi anemia pathway by patient-derived FANCA mutants. *Hum Mol Genet*, 11, 3125-34.
- ADESSI, C., MATTON, G., AYALA, G., TURCATTI, G., MERMOD, J. J., MAYER, P. & KAWASHIMA, E. 2000. Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res*, 28, E87.
- ADEY, A., MORRISON, H. G., ASAN, XUN, X., KITZMAN, J. O., TURNER, E. H., STACKHOUSE, B., MACKENZIE, A. P., CARUCCIO, N. C., ZHANG, X. & SHENDURE, J. 2010. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol*, 11, R119.
- ADZHUBEI, I. A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V. E., GERASIMOVA, A., BORK, P., KONDRASHOV, A. S. & SUNYAEV, S. R. 2010. A method and server for predicting damaging missense mutations. *Nat Methods*, 7, 248-9.
- AHMAD, A., ENZLIN, J. H., BHAGWAT, N. R., WIJGERS, N., RAAMS, A., APPELDOORN, E., THEIL, A. F., JH, J. H., VERMEULEN, W., NG, J. J., SCHARER, O. D. & NIEDERNHOFER, L. J. 2010. Mislocalization of XPF-ERCC1 nuclease contributes to reduced DNA repair in XP-F patients. *PLoS Genet*, 6, e1000871.
- AHMAD, A., ROBINSON, A. R., DUENSING, A., VAN DRUNEN, E., BEVERLOO, H. B., WEISBERG, D. B., HASTY, P., HOEIJMAKERS, J. H. & NIEDERNHOFER, L. J. 2008. ERCC1-XPF endonuclease facilitates DNA double-strand break repair. *Mol Cell Biol*, 28, 5082-92.
- AL-MINAWI, A. Z., LEE, Y. F., HAKANSSON, D., JOHANSSON, F., LUNDIN, C., SALEH-GOHARI, N., SCHULTZ, N., JENSSEN, D., BRYANT, H. E., MEUTH, M., HINZ, J. M. & HELLEDAY, T. 2009. The ERCC1/XPF endonuclease is required for completion of homologous recombination at DNA replication forks stalled by inter-strand cross-links. *Nucleic Acids Res*, 37, 6400-13.
- ALBERT, T. J., MOLLA, M. N., MUZNY, D. M., NAZARETH, L., WHEELER, D., SONG, X., RICHMOND, T. A., MIDDLE, C. M., RODESCH, M. J., PACKARD, C. J., WEINSTOCK, G. M. & GIBBS, R. A. 2007. Direct selection of human genomic loci by microarray hybridization. *Nat Methods*, 4, 903-5.
- ALMARZA, E., RIO, P., MEZA, N. W., ALDEA, M., AGIRRE, X., GUENECHEA, G., SEGOVIA, J. C. & BUEREN, J. A. 2007. Characteristics of lentiviral vectors harboring the proximal promoter of the *vav* proto-oncogene: a weak and efficient promoter for gene therapy. *Mol Ther*, 15, 1487-94.
- ALPI, A. F., PACE, P. E., BABU, M. M. & PATEL, K. J. 2008. Mechanistic insight into site-restricted monoubiquitination of FANCD2 by Ube2t, FANCL, and FANCI. *Mol Cell*, 32, 767-77.
- ALTER, B. P. 2014. Fanconi anemia and the development of leukemia. *Best Pract Res Clin Haematol*, 27, 214-21.
- ALTER, B. P., GREENE, M. H., VELAZQUEZ, I. & ROSENBERG, P. S. 2003. Cancer in Fanconi anemia. *Blood*, 101, 2072.

- ALTER, B. P., JOENJE, H., OOSTRA, A. B. & PALS, G. 2005. Fanconi anemia: adult head and neck cancer and hematopoietic mosaicism. *Arch Otolaryngol Head Neck Surg*, 131, 635-9.
- ALTER, B. P., ROSENBERG, P. S. & BRODY, L. C. 2007. Clinical and molecular features associated with biallelic mutations in FANCD1/BRCA2. *J Med Genet*, 44, 1-9.
- AMEZIANE, N., ERRAMI, A., LEVEILLE, F., FONTAINE, C., DE VRIES, Y., VAN SPAENDONK, R. M., DE WINTER, J. P., PALS, G. & JOENJE, H. 2008. Genetic subtyping of Fanconi anemia by comprehensive mutation screening. *Hum Mutat*, 29, 159-66.
- AMEZIANE, N., SIE, D., DENTRO, S., ARIYUREK, Y., KERKHOVEN, L., JOENJE, H., DORSMAN, J. C., YLSTRA, B., GILLE, J. J., SISTERMANS, E. A. & DE WINTER, J. P. 2012. Diagnosis of fanconi anemia: mutation analysis by next-generation sequencing. *Anemia*, 2012, 132856.
- ANDERSEN, S. L., BERGSTALH, D. T., KOHL, K. P., LAROCQUE, J. R., MOORE, C. B. & SEKELSKY, J. 2009. Drosophila MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination. *Mol Cell*, 35, 128-35.
- ANDERSSON, B. S., SADEGHI, T., SICILIANO, M. J., LEGERSKI, R. & MURRAY, D. 1996. Nucleotide excision repair genes as determinants of cellular sensitivity to cyclophosphamide analogs. *Cancer Chemother Pharmacol*, 38, 406-16.
- AUERBACH, A. D. 1988. A test for Fanconi's anemia. *Blood*, 72, 366-7.
- AUERBACH, A. D. 1994. Umbilical cord blood transplants for genetic disease: diagnostic and ethical issues in fetal studies. *Blood Cells*, 20, 303-9.
- AUERBACH, A. D. 2003. Diagnosis of fanconi anemia by diepoxybutane analysis. *Curr Protoc Hum Genet*, Chapter 8, Unit 8 7.
- AUERBACH, A. D. 2009. Fanconi anemia and its diagnosis. *Mutat Res*, 668, 4-10.
- AZQUETA, A. & COLLINS, A. R. 2013. The essential comet assay: a comprehensive guide to measuring DNA damage and repair. *Arch Toxicol*, 87, 949-68.
- BAKKER, J. L., VAN MIL, S. E., CROSSAN, G., SABBAGHIAN, N., DE LEENEER, K., POPPE, B., ADANK, M., GILLE, H., VERHEUL, H., MEIJERS-HEIJBOER, H., DE WINTER, J. P., CLAES, K., TISCHKOWITZ, M. & WAISFISZ, Q. 2012. Analysis of the Novel Fanconi Anemia Gene SLX4/FANCP in Familial Breast Cancer Cases. *Hum Mutat*.
- BAKKER, J. L., VAN MIL, S. E., CROSSAN, G., SABBAGHIAN, N., DE LEENEER, K., POPPE, B., ADANK, M., GILLE, H., VERHEUL, H., MEIJERS-HEIJBOER, H., DE WINTER, J. P., CLAES, K., TISCHKOWITZ, M. & WAISFISZ, Q. 2013. Analysis of the novel fanconi anemia gene SLX4/FANCP in familial breast cancer cases. *Hum Mutat*, 34, 70-3.
- BARALLE, D. & BARALLE, M. 2005. Splicing in action: assessing disease causing sequence changes. *J Med Genet*, 42, 737-48.
- BENSON, F. E., BAUMANN, P. & WEST, S. C. 1998. Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. *Nature*, 391, 401-4.
- BESSHO, T., SANCAR, A., THOMPSON, L. H. & THELEN, M. P. 1997. Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex. *J Biol Chem*, 272, 3833-7.
- BIANCO, P. R., TRACY, R. B. & KOWALCZYKOWSKI, S. C. 1998. DNA strand exchange proteins: a biochemical and physical comparison. *Front Biosci*, 3, D570-603.
- BOGLIOLO, M., LYAKHOVICH, A., CALLEN, E., CASTELLA, M., CAPPELLI, E., RAMIREZ, M. J., CREUS, A., MARCOS, R., KALB, R., NEVELING, K., SCHINDLER, D. & SURRELLES, J. 2007. Histone H2AX and Fanconi anemia FANCD2 function in the same pathway to maintain chromosome stability. *EMBO J*, 26, 1340-51.
- BOGLIOLO, M., SCHUSTER, B., STOEPKER, C., DERKUNT, B., SU, Y., RAAMS, A., TRUJILLO, J. P., MINGUILLON, J., RAMIREZ, M. J., PUJOL, R., CASADO, J. A., BANOS, R., RIO, P., KNIES, K., ZUNIGA, S., BENITEZ, J., BUEREN, J. A., JASPERS, N. G., SCHARER, O. D., DE WINTER, J. P.,

- SCHINDLER, D. & SURRELLES, J. 2013. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet*, 92, 800-6.
- BOGLIOLO, M. & SURRELLES, J. 2015. Fanconi anemia: a model disease for studies on human genetics and advanced therapeutics. *Curr Opin Genet Dev*, 33, 32-40.
- BONACHE, S., DE LA HOYA, M., GUTIERREZ-ENRIQUEZ, S., TENES, A., MASAS, M., BALMANA, J. & DIEZ, O. 2013. Mutation analysis of the SHFM1 gene in breast/ovarian cancer families. *J Cancer Res Clin Oncol*, 139, 529-32.
- BOUCLAKA, C., ABDELHAK, S., AMOURI, A., BEN ABID, H., HADIJI, S., FRIKHA, M., BEN OTHMAN, T., AMRI, F., AYADI, H., HACHICHA, M., REBAI, A., SAAD, A. & DELLAGI, K. 2003. Fanconi anemia in Tunisia: high prevalence of group A and identification of new FANCA mutations. *J Hum Genet*, 48, 352-61.
- BRIDGE, W. L., VANDENBERG, C. J., FRANKLIN, R. J. & HIOM, K. 2005. The BRIP1 helicase functions independently of BRCA1 in the Fanconi anemia pathway for DNA crosslink repair. *Nat Genet*, 37, 953-7.
- BUNTING, S. F., CALLEN, E., WONG, N., CHEN, H. T., POLATO, F., GUNN, A., BOTHMER, A., FELDHahn, N., FERNANDEZ-CAPETILLO, O., CAO, L., XU, X., DENG, C. X., FINKEL, T., NUSSENZWEIG, M., STARK, J. M. & NUSSENZWEIG, A. 2010. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell*, 141, 243-54.
- BURROWS, A. E. & ELLEDGE, S. J. 2008. How ATR turns on: TopBP1 goes on ATRIP with ATR. *Genes Dev*, 22, 1416-21.
- BUTTURINI, A., GALE, R. P., VERLANDER, P. C., ADLER-BRECHER, B., GILLIO, A. P. & AUERBACH, A. D. 1994. Hematologic abnormalities in Fanconi anemia: an International Fanconi Anemia Registry study. *Blood*, 84, 1650-5.
- CALLEN, E., CASADO, J. A., TISCHKOWITZ, M. D., BUEREN, J. A., CREUS, A., MARCOS, R., DASÍ, A., ESTELLA, J. M., MUNOZ, A., ORTEGA, J. J., DE WINTER, J., JOENJE, H., SCHINDLER, D., HANENBERG, H., HODGSON, S. V., MATHEW, C. G. & SURRELLES, J. 2005. A common founder mutation in FANCA underlies the world's highest prevalence of Fanconi anemia in Gypsy families from Spain. *Blood*, 105, 1946-9.
- CALLEN, E., TISCHKOWITZ, M. D., CREUS, A., MARCOS, R., BUEREN, J. A., CASADO, J. A., MATHEW, C. G. & SURRELLES, J. 2004. Quantitative PCR analysis reveals a high incidence of large intragenic deletions in the FANCA gene in Spanish Fanconi anemia patients. *Cytogenet Genome Res*, 104, 341-5.
- CANTOR, S. B. & ANDREASSEN, P. R. 2006. Assessing the link between BACH1 and BRCA1 in the FA pathway. *Cell Cycle*, 5, 164-7.
- CASADO, A. J., CALLEN, E., JACOME, A., RIO, P., CASTELLA, M., LOBITZ, S., FERRO, T., MUNOZ, A., SEVILLA, J., CANTALEJO, A., CELA, E., CERVERA, J., SANCHEZ-CALERO, J., BADELL, I., ESTELLA, J., DASÍ, A., OLIVE, T., JOSE ORTEGA, J., RODRIGUEZ-VILLA, A., TAPIA, M., MOLINES, A., MADERO, L., SEGOVIA, J. C., NEVELING, K., KALB, R., SCHINDLER, D., HANENBERG, H., SURRELLES, J. & BUEREN, J. A. 2007. A comprehensive strategy for the subtyping of patients with Fanconi anaemia: conclusions from the Spanish Fanconi Anemia Research Network. *J Med Genet*, 44, 241-9.
- CASTELLA, M., PUJOL, R., CALLEN, E., RAMIREZ, M. J., CASADO, J. A., TALAVERA, M., FERRO, T., MUNOZ, A., SEVILLA, J., MADERO, L., CELA, E., BELENDEZ, C., DE HEREDIA, C. D., OLIVE, T., DE TOLEDO, J. S., BADELL, I., ESTELLA, J., DASÍ, A., RODRIGUEZ-VILLA, A., GOMEZ, P., TAPIA, M., MOLINES, A., FIGUERA, A., BUEREN, J. A. & SURRELLES, J. 2011a. Chromosome fragility in patients with Fanconi anaemia: diagnostic implications and clinical impact. *J Med Genet*, 48, 242-50.

- CASTELLA, M., PUJOL, R., CALLEN, E., TRUJILLO, J. P., CASADO, J. A., GILLE, H., LACH, F. P., AUERBACH, A. D., SCHINDLER, D., BENITEZ, J., PORTO, B., FERRO, T., MUNOZ, A., SEVILLA, J., MADERO, L., CELA, E., BELENDEZ, C., DE HEREDIA, C. D., OLIVE, T., DE TOLEDO, J. S., BADELL, I., TORRENT, M., ESTELLA, J., DASÍ, A., RODRIGUEZ-VILLA, A., GOMEZ, P., BARBOT, J., TAPIA, M., MOLINES, A., FIGUERA, A., BUEREN, J. A. & SURRALLES, J. 2011b. Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. *Blood*, 117, 3759-69.
- CASTILLO BOSCH, P., SEGURA-BAYONA, S., KOOLE, W., VAN HETEREN, J. T., DEWAR, J. M., TIJSTERMAN, M. & KNIPSCHER, P. 2014. FANCI promotes DNA synthesis through G-quadruplex structures. *EMBO J*, 33, 2521-33.
- CATUCCI, I., COLOMBO, M., VERDERIO, P., BERNARD, L., FICARAZZI, F., MARIETTE, F., BARILE, M., PEISSEL, B., CATTANEO, E., MANOUKIAN, S., RADICE, P. & PETERLONGO, P. 2012. Sequencing analysis of SLX4/FANCP gene in Italian familial breast cancer cases. *PLoS One*, 7, e31038.
- CICCIA, A., MCDONALD, N. & WEST, S. C. 2008. Structural and functional relationships of the XPF/MUS81 family of proteins. *Annu Rev Biochem*, 77, 259-87.
- CLEAVER, J. E. 1968. Defective repair replication of DNA in xeroderma pigmentosum. *Nature*, 218, 652-6.
- CLEAVER, J. E., THOMPSON, L. H., RICHARDSON, A. S. & STATES, J. C. 1999. A summary of mutations in the UV-sensitive disorders: xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. *Hum Mutat*, 14, 9-22.
- CLINGEN, P. H., DE SILVA, I. U., MCHUGH, P. J., GHADDESSY, F. J., TILBY, M. J., THURSTON, D. E. & HARTLEY, J. A. 2005. The XPF-ERCC1 endonuclease and homologous recombination contribute to the repair of minor groove DNA interstrand crosslinks in mammalian cells produced by the pyrrolo[2,1-c][1,4]benzodiazepine dimer SJG-136. *Nucleic Acids Res*, 33, 3283-91.
- COHN, M. A., KOWAL, P., YANG, K., HAAS, W., HUANG, T. T., GYGI, S. P. & D'ANDREA, A. D. 2007. A UAF1-containing multisubunit protein complex regulates the Fanconi anemia pathway. *Mol Cell*, 28, 786-97.
- CONSTANTINO, A., TARSOUNAS, M., KAROW, J. K., BROSH, R. M., BOHR, V. A., HICKSON, I. D. & WEST, S. C. 2000. Werner's syndrome protein (WRN) migrates Holliday junctions and co-localizes with RPA upon replication arrest. *EMBO Rep*, 1, 80-4.
- COX, P. M., GIBSON, R. A., MORGAN, N. & BRUETON, L. A. 1997. VACTERL with hydrocephalus in twins due to Fanconi anemia (FA): mutation in the FAC gene. *Am J Med Genet*, 68, 86-90.
- CRABB, D. W., EDENBERG, H. J., BOSRON, W. F. & LI, T. K. 1989. Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant. *J Clin Invest*, 83, 314-6.
- CROSSAN, G. P. & PATEL, K. J. 2012. The Fanconi anaemia pathway orchestrates incisions at sites of crosslinked DNA. *J Pathol*, 226, 326-37.
- CROSSAN, G. P., VAN DER WEYDEN, L., ROSADO, I. V., LANGEVIN, F., GAILLARD, P. H., MCINTYRE, R. E., GALLAGHER, F., KETTUNEN, M. I., LEWIS, D. Y., BRINDLE, K., ARENDS, M. J., ADAMS, D. J. & PATEL, K. J. 2011. Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. *Nat Genet*, 43, 147-52.
- CYBULSKI, K. E. & HOWLETT, N. G. 2011. FANCP/SLX4: a Swiss army knife of DNA interstrand crosslink repair. *Cell Cycle*, 10, 1757-63.
- CHANDRASEKHARAPPA, S. C., LACH, F. P., KIMBLE, D. C., KAMAT, A., TEER, J. K., DONOVAN, F. X., FLYNN, E., SEN, S. K., THONGTHIP, S., SANBORN, E., SMOGORZEWSKA, A., AUERBACH, A. D. & OSTRANDER, E. A. 2013. Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia. *Blood*, 121, e138-48.

- CHAUDHURY, I., STROIK, D. R. & SOBECK, A. 2014. FANCD2-controlled chromatin access of the Fanconi-associated nuclease FAN1 is crucial for the recovery of stalled replication forks. *Mol Cell Biol*, 34, 3939-54.
- CHEN, C. H., FERREIRA, J. C., GROSS, E. R. & MOCHLY-ROSEN, D. 2014. Targeting aldehyde dehydrogenase 2: new therapeutic opportunities. *Physiol Rev*, 94, 1-34.
- CHENG, N. C., VAN DE VRUGT, H. J., VAN DER VALK, M. A., OOSTRA, A. B., KRIMPENFORT, P., DE VRIES, Y., JOENJE, H., BERNS, A. & ARWERT, F. 2000. Mice with a targeted disruption of the Fanconi anemia homolog Fanca. *Hum Mol Genet*, 9, 1805-11.
- CHOI, M., SCHOLL, U. I., JI, W., LIU, T., TIKHONOVA, I. R., ZUMBO, P., NAYIR, A., BAKKALOGLU, A., OZEN, S., SANJAD, S., NELSON-WILLIAMS, C., FARHI, A., MANE, S. & LIFTON, R. P. 2009. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A*, 106, 19096-101.
- D'ANDREA, A. D. 2010. Susceptibility pathways in Fanconi's anemia and breast cancer. *N Engl J Med*, 362, 1909-19.
- D'ANDREA, A. D. & GROMPE, M. 1997. Molecular biology of Fanconi anemia: implications for diagnosis and therapy. *Blood*, 90, 1725-36.
- DAMIA, G., IMPERATORI, L., STEFANINI, M. & D'INCALCI, M. 1996. Sensitivity of CHO mutant cell lines with specific defects in nucleotide excision repair to different anti-cancer agents. *Int J Cancer*, 66, 779-83.
- DE GARIBAY, G. R., DIAZ, A., GAVINA, B., ROMERO, A., GARRE, P., VEGA, A., BLANCO, A., TOSAR, A., DIEZ, O., PEREZ-SEGURA, P., DIAZ-RUBIO, E., CALDES, T. & DE LA HOYA, M. 2013. Low prevalence of SLX4 loss-of-function mutations in non-BRCA1/2 breast and/or ovarian cancer families. *Eur J Hum Genet*, 21, 883-6.
- DE LAAT, W. L., SIJBERS, A. M., ODIJK, H., JASPERS, N. G. & HOEIJMAKERS, J. H. 1998. Mapping of interaction domains between human repair proteins ERCC1 and XPF. *Nucleic Acids Res*, 26, 4146-52.
- DE SILVA, I. U., MCHUGH, P. J., CLINGEN, P. H. & HARTLEY, J. A. 2000. Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. *Mol Cell Biol*, 20, 7980-90.
- DE SILVA, I. U., MCHUGH, P. J., CLINGEN, P. H. & HARTLEY, J. A. 2002. Defects in interstrand cross-link uncoupling do not account for the extreme sensitivity of ERCC1 and XPF cells to cisplatin. *Nucleic Acids Res*, 30, 3848-56.
- DE WINTER, J. P. & JOENJE, H. 2009. The genetic and molecular basis of Fanconi anemia. *Mutat Res*, 668, 11-9.
- DE WINTER, J. P., LEVEILLE, F., VAN BERKEL, C. G., ROOIMANS, M. A., VAN DER WEEL, L., STELTENPOOL, J., DEMUTH, I., MORGAN, N. V., ALON, N., BOSNOYAN-COLLINS, L., LIGHTFOOT, J., LEEGWATER, P. A., WAISFISZ, Q., KOMATSU, K., ARWERT, F., PRONK, J. C., MATHEW, C. G., DIGWEED, M., BUCHWALD, M. & JOENJE, H. 2000. Isolation of a cDNA representing the Fanconi anemia complementation group E gene. *Am J Hum Genet*, 67, 1306-8.
- DE WINTER, J. P., WAISFISZ, Q., ROOIMANS, M. A., VAN BERKEL, C. G., BOSNOYAN-COLLINS, L., ALON, N., CARREAU, M., BENDER, O., DEMUTH, I., SCHINDLER, D., PRONK, J. C., ARWERT, F., HOEHN, H., DIGWEED, M., BUCHWALD, M. & JOENJE, H. 1998. The Fanconi anaemia group G gene FANCG is identical with XRCC9. *Nat Genet*, 20, 281-3.
- DEANS, A. J. & WEST, S. C. 2009. FANCM connects the genome instability disorders Bloom's Syndrome and Fanconi Anemia. *Mol Cell*, 36, 943-53.
- DEANS, A. J. & WEST, S. C. 2011. DNA interstrand crosslink repair and cancer. *Nat Rev Cancer*, 11, 467-80.

- DIKIC, I. & ROBERTSON, M. 2012. Ubiquitin ligases and beyond. *BMC Biol*, 10, 22.
- DOHM, J. C., LOTTAZ, C., BORODINA, T. & HIMMELBAUER, H. 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res*, 36, e105.
- DOMCZEK, S. M., TANG, J., STOPFER, J., LILLI, D. R., HAMEL, N., TISCHKOWITZ, M., MONTEIRO, A. N., MESSICK, T. E., POWERS, J., YONKER, A., COUCH, F. J., GOLDGAR, D. E., DAVIDSON, H. R., NATHANSON, K. L., FOULKES, W. D. & GREENBERG, R. A. 2013. Biallelic deleterious BRCA1 mutations in a woman with early-onset ovarian cancer. *Cancer Discov*, 3, 399-405.
- DORSMAN, J. C., LEVITUS, M., ROCKX, D., ROOIMANS, M. A., OOSTRA, A. B., HAITJEMA, A., BAKKER, S. T., STELTENPOOL, J., SCHULER, D., MOHAN, S., SCHINDLER, D., ARWERT, F., PALS, G., MATHEW, C. G., WAISFISZ, Q., DE WINTER, J. P. & JOENJE, H. 2007. Identification of the Fanconi anemia complementation group I gene, FANCI. *Cell Oncol*, 29, 211-8.
- DUTRILLAUX, B., AURIAS, A., DUTRILLAUX, A. M., BURIOT, D. & PRIEUR, M. 1982. The cell cycle of lymphocytes in Fanconi anemia. *Hum Genet*, 62, 327-32.
- EID, J., FEHR, A., GRAY, J., LUONG, K., LYLE, J., OTTO, G., PELUSO, P., RANK, D., BAYBAYAN, P., BETTMAN, B., BIBILLO, A., BJORNSON, K., CHAUDHURI, B., CHRISTIANS, F., CICERO, R., CLARK, S., DALAL, R., DEWINTER, A., DIXON, J., FOQUET, M., GAERTNER, A., HARDENBOL, P., HEINER, C., HESTER, K., HOLDEN, D., KEARNS, G., KONG, X., KUSE, R., LACROIX, Y., LIN, S., LUNDQUIST, P., MA, C., MARKS, P., MAXHAM, M., MURPHY, D., PARK, I., PHAM, T., PHILLIPS, M., ROY, J., SEBRA, R., SHEN, G., SORENSON, J., TOMANEY, A., TRAVERS, K., TRULSON, M., VIECELI, J., WEGENER, J., WU, D., YANG, A., ZACCARIN, D., ZHAO, P., ZHONG, F., KORLACH, J. & TURNER, S. 2009. Real-time DNA sequencing from single polymerase molecules. *Science*, 323, 133-8.
- ENZLIN, J. H. & SCHARER, O. D. 2002. The active site of the DNA repair endonuclease XPF-ERCC1 forms a highly conserved nuclease motif. *EMBO J*, 21, 2045-53.
- EVANS, E., MOGGS, J. G., HWANG, J. R., EGLY, J. M. & WOOD, R. D. 1997. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J*, 16, 6559-73.
- FANCONI ANAEMIA/BREAST CANCER, C. 1996. Positional cloning of the Fanconi anaemia group A gene. *Nat Genet*, 14, 324-8.
- FANCONI, G. 1927. Familiäre, infantile perniciosähnliche Anämie (perniziöses Blutbild und Konstitution). *Jahrbuch für Kinderheilkunde und physische Erziehung*, 117, 257-280.
- FEKAIRI, S., SCAGLIONE, S., CHAHWAN, C., TAYLOR, E. R., TISSIER, A., COULON, S., DONG, M. Q., RUSE, C., YATES, J. R., 3RD, RUSSELL, P., FUCHS, R. P., MCGOWAN, C. H. & GAILLARD, P. H. 2009. Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell*, 138, 78-89.
- FERNANDEZ-RODRIGUEZ, J., QUILES, F., BLANCO, I., TEULE, A., FELIUBADALO, L., VALLE, J. D., SALINAS, M., IZQUIERDO, A., DARDER, E., SCHINDLER, D., CAPELLA, G., BRUNET, J., LAZARO, C. & PUJANA, M. A. 2012. Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families. *BMC Cancer*, 12, 84.
- FISHER, L. A., BESSHO, M. & BESSHO, T. 2008. Processing of a psoralen DNA interstrand cross-link by XPF-ERCC1 complex in vitro. *J Biol Chem*, 283, 1275-81.
- FRIEDBERG, E. C., AGUILERA, A., GELLERT, M., HANAWALT, P. C., HAYS, J. B., LEHMANN, A. R., LINDAHL, T., LOWNDES, N., SARASIN, A. & WOOD, R. D. 2006a. DNA repair: from molecular mechanism to human disease. *DNA Repair (Amst)*, 5, 986-96.
- FRIEDBERG, E. C., WALKER, G. C., SIEDE, W., WOOD, R. D., SCHULTZ, R. A. & ELLENBERGER, T. 2006b. DNA Repair and Mutagenesis, 2nd Ed. *ASM Press, Washington, D. C.*
- FRISAN, T., LEVITSKY, V. & MASUCCI, M. 2001. Generation of Lymphoblastoid Cell Lines (LCLs). In: WILSON, J. & MAY, G. W. (eds.) *Epstein-Barr Virus Protocols*. Humana Press.

- GAILLARD, P. H. & WOOD, R. D. 2001. Activity of individual ERCC1 and XPF subunits in DNA nucleotide excision repair. *Nucleic Acids Res*, 29, 872-9.
- GARAYCOECHEA, J. I., CROSSAN, G. P., LANGEVIN, F., DALY, M., ARENDS, M. J. & PATEL, K. J. 2012. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature*, 489, 571-5.
- GARCIA-HIGUERA, I., KUANG, Y., DENHAM, J. & D'ANDREA, A. D. 2000. The fanconi anemia proteins FANCA and FANCG stabilize each other and promote the nuclear accumulation of the Fanconi anemia complex. *Blood*, 96, 3224-30.
- GARCIA-HIGUERA, I., KUANG, Y., NAF, D., WASIK, J. & D'ANDREA, A. D. 1999. Fanconi anemia proteins FANCA, FANCC, and FANCG/XRCC9 interact in a functional nuclear complex. *Mol Cell Biol*, 19, 4866-73.
- GARCIA-HIGUERA, I., TANIGUCHI, T., GANESAN, S., MEYN, M. S., TIMMERS, C., HEJNA, J., GROMPE, M. & D'ANDREA, A. D. 2001. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell*, 7, 249-62.
- GARNER, E. & SMOGORZEWSKA, A. 2011. Ubiquitylation and the Fanconi anemia pathway. *FEBS Lett*, 585, 2853-60.
- GENOVESE, P., SCHIROLI, G., ESCOBAR, G., DI TOMASO, T., FIRRITO, C., CALABRIA, A., MOI, D., MAZZIERI, R., BONINI, C., HOLMES, M. C., GREGORY, P. D., VAN DER BURG, M., GENTNER, B., MONTINI, E., LOMBARDO, A. & NALDINI, L. 2014. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature*, 510, 235-40.
- GILLE, J. J., FLOOR, K., KERKHOVEN, L., AMEZIANE, N., JOENJE, H. & DE WINTER, J. P. 2012. Diagnosis of Fanconi Anemia: Mutation Analysis by Multiplex Ligation-Dependent Probe Amplification and PCR-Based Sanger Sequencing. *Anemia*, 2012, 603253.
- GLICKMAN, M. H. & CIECHANOVER, A. 2002. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev*, 82, 373-428.
- GNIRKE, A., MELNIKOV, A., MAGUIRE, J., ROGOV, P., LEPROUST, E. M., BROCKMAN, W., FENNEL, T., GIANNOUKOS, G., FISHER, S., RUSS, C., GABRIEL, S., JAFFE, D. B., LANDER, E. S. & NUSBAUM, C. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol*, 27, 182-9.
- GONZALEZ-PEREZ, A. & LOPEZ-BIGAS, N. 2011. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, *Condel*. *Am J Hum Genet*, 88, 440-9.
- GRAVEL, S., CHAPMAN, J. R., MAGILL, C. & JACKSON, S. P. 2008. DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes Dev*, 22, 2767-72.
- GREAVES, M. F. & WIEMELS, J. 2003. Origins of chromosome translocations in childhood leukaemia. *Nat Rev Cancer*, 3, 639-49.
- GREENBERG, R. A., SOBHIAN, B., PATHANIA, S., CANTOR, S. B., NAKATANI, Y. & LIVINGSTON, D. M. 2006. Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes. *Genes Dev*, 20, 34-46.
- GREGG, S. Q., ROBINSON, A. R. & NIEDERNHOFER, L. J. 2011. Physiological consequences of defects in ERCC1-XPF DNA repair endonuclease. *DNA Repair (Amst)*, 10, 781-91.
- GRESNER, P., GROMADZINSKA, J., TWARDOWSKA, E., RYDZYNSKI, K. & WASOWICZ, W. 2014. Rad51C: a novel suppressor gene modulates the risk of head and neck cancer. *Mutat Res Fundam Mol Mech Mutagen*, 762, 47-54.
- GROMPE, M. & D'ANDREA, A. 2003. Reply to "Involvement of oxidative stress in Fanconi's anaemia: from phenotype to FA protein functions". *Nat Rev Cancer*, 3.

- GROSS, M., HANENBERG, H., LOBITZ, S., FRIEDL, R., HERTERICH, S., DIETRICH, R., GRUHN, B., SCHINDLER, D. & HOEHN, H. 2002. Reverse mosaicism in Fanconi anemia: natural gene therapy via molecular self-correction. *Cytogenet Genome Res*, 98, 126-35.
- HANENBERG, H., BATISH, S. D., POLLOK, K. E., VIETEN, L., VERLANDER, P. C., LEURS, C., COOPER, R. J., GOTTSCHKE, K., HANELINE, L., CLAPP, D. W., LOBITZ, S., WILLIAMS, D. A. & AUERBACH, A. D. 2002. Phenotypic correction of primary Fanconi anemia T cells with retroviral vectors as a diagnostic tool. *Exp Hematol*, 30, 410-20.
- HEBRA, F. K. M. 1874. On diseases of the skin, including exanthemata. *New Sydenham Soc*, 61, 252-8.
- HEJNA, J. A., TIMMERS, C. D., REIFSTECK, C., BRUUN, D. A., LUCAS, L. W., JAKOBS, P. M., TOTH-FEJEL, S., UNSWORTH, N., CLEMENS, S. L., GARCIA, D. K., NAYLOR, S. L., THAYER, M. J., OLSON, S. B., GROMPE, M. & MOSES, R. E. 2000. Localization of the Fanconi anemia complementation group D gene to a 200-kb region on chromosome 3p25.3. *Am J Hum Genet*, 66, 1540-51.
- HELLEDAY, T., LO, J., VAN GENT, D. C. & ENGELWARD, B. P. 2007. DNA double-strand break repair: from mechanistic understanding to cancer treatment. *DNA Repair (Amst)*, 6, 923-35.
- HIRA, A., YABE, H., YOSHIDA, K., OKUNO, Y., SHIRAIISHI, Y., CHIBA, K., TANAKA, H., MIYANO, S., NAKAMURA, J., KOJIMA, S., OGAWA, S., MATSUO, K., TAKATA, M. & YABE, M. 2013. Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients. *Blood*, 122, 3206-9.
- HIRA, A., YOSHIDA, K., SATO, K., OKUNO, Y., SHIRAIISHI, Y., CHIBA, K., TANAKA, H., MIYANO, S., SHIMAMOTO, A., TAHARA, H., ITO, E., KOJIMA, S., KURUMIZAKA, H., OGAWA, S., TAKATA, M., YABE, H. & YABE, M. 2015. Mutations in the Gene Encoding the E2 Conjugating Enzyme UBE2T Cause Fanconi Anemia. *Am J Hum Genet*, 96, 1001-7.
- HIRAI, Y., KODAMA, Y., MORIWAKI, S., NODA, A., CULLINGS, H. M., MACPHEE, D. G., KODAMA, K., MABUCHI, K., KRAEMER, K. H., LAND, C. E. & NAKAMURA, N. 2006. Heterozygous individuals bearing a founder mutation in the XPA DNA repair gene comprise nearly 1% of the Japanese population. *Mutat Res*, 601, 171-8.
- HIRSCHHORN, R., YANG, D. R., PUCK, J. M., HUIE, M. L., JIANG, C. K. & KURLANDSKY, L. E. 1996. Spontaneous in vivo reversion to normal of an inherited mutation in a patient with adenosine deaminase deficiency. *Nat Genet*, 13, 290-5.
- HO, G. P., MARGOSSIAN, S., TANIGUCHI, T. & D'ANDREA, A. D. 2006. Phosphorylation of FANCD2 on two novel sites is required for mitomycin C resistance. *Mol Cell Biol*, 26, 7005-15.
- HODGES, E., XUAN, Z., BALIJA, V., KRAMER, M., MOLLA, M. N., SMITH, S. W., MIDDLE, C. M., RODESCH, M. J., ALBERT, T. J., HANNON, G. J. & MCCOMBIE, W. R. 2007. Genome-wide in situ exon capture for selective resequencing. *Nat Genet*, 39, 1522-7.
- HOFMANN, K. 2009. Ubiquitin-binding domains and their role in the DNA damage response. *DNA Repair (Amst)*, 8, 544-556.
- HOLDEN, S. T., COX, J. J., KESTERTON, I., THOMAS, N. S., CARR, C. & WOODS, C. G. 2006. Fanconi anaemia complementation group B presenting as X linked VACTERL with hydrocephalus syndrome. *J Med Genet*, 43, 750-4.
- HOLLOWAY, J. K., MOHAN, S., BALMUS, G., SUN, X., MODZELEWSKI, A., BORST, P. L., FREIRE, R., WEISS, R. S. & COHEN, P. E. 2011. Mammalian BTBD12 (SLX4) protects against genomic instability during mammalian spermatogenesis. *PLoS Genet*, 7, e1002094.
- HOWLETT, N. G., HARNEY, J. A., REGO, M. A., KOLLING, F. W. T. & GLOVER, T. W. 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, 28935-42.

- HOWLETT, N. G., TANIGUCHI, T., OLSON, S., COX, B., WAISFISZ, Q., DE DIE-SMULDERS, C., PERSKY, N., GROMPE, M., JOENJE, H., PALS, G., IKEDA, H., FOX, E. A. & D'ANDREA, A. D. 2002. Biallelic inactivation of BRCA2 in Fanconi anemia. *Science*, 297, 606-9.
- HUANG, T. T., NIJMAN, S. M., MIRCHANDANI, K. D., GALARDY, P. J., COHN, M. A., HAAS, W., GYGI, S. P., PLOEGH, H. L., BERNARDS, R. & D'ANDREA, A. D. 2006. Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat Cell Biol*, 8, 339-47.
- IP, S. C., RASS, U., BLANCO, M. G., FLYNN, H. R., SKEHEL, J. M. & WEST, S. C. 2008. Identification of Holliday junction resolvases from humans and yeast. *Nature*, 456, 357-61.
- ISHIAI, M., KITAO, H., SMOGORZEWSKA, A., TOMIDA, J., KINOMURA, A., UCHIDA, E., SABERI, A., KINOSHITA, E., KINOSHITA-KIKUTA, E., KOIKE, T., TASHIRO, S., ELLEDGE, S. J. & TAKATA, M. 2008. FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nat Struct Mol Biol*, 15, 1138-46.
- JAMUAR, S. S., LAM, A. T., KIRCHER, M., D'GAMA, A. M., WANG, J., BARRY, B. J., ZHANG, X., HILL, R. S., PARTLOW, J. N., ROZZO, A., SERVATTALAB, S., MEHTA, B. K., TOPCU, M., AMROM, D., ANDERMANN, E., DAN, B., PARRINI, E., GUERRINI, R., SCHEFFER, I. E., BERKOVIC, S. F., LEVENTER, R. J., SHEN, Y., WU, B. L., BARKOVICH, A. J., SAHIN, M., CHANG, B. S., BAMSHAD, M., NICKERSON, D. A., SHENDURE, J., PODURI, A., YU, T. W. & WALSH, C. A. 2014. Somatic mutations in cerebral cortical malformations. *N Engl J Med*, 371, 733-43.
- JANG, S. K., KRAUSSLICH, H. G., NICKLIN, M. J., DUKE, G. M., PALMENBERG, A. C. & WIMMER, E. 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol*, 62, 2636-43.
- JASPERS, N. G. & BOOTSMA, D. 1982. Genetic heterogeneity in ataxia-telangiectasia studied by cell fusion. *Proc Natl Acad Sci U S A*, 79, 2641-4.
- JASPERS, N. G., RAAMS, A., SILENGO, M. C., WIJGERS, N., NIEDERNHOFER, L. J., ROBINSON, A. R., GIGLIA-MARI, G., HOOGSTRATEN, D., KLEIJER, W. J., HOEIJMAKERS, J. H. & VERMEULEN, W. 2007. First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. *Am J Hum Genet*, 80, 457-66.
- JENSEN, R. B., CARREIRA, A. & KOWALCZYKOWSKI, S. C. 2010. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature*, 467, 678-83.
- JOENJE, H., LO TEN FOE, J. R., OOSTRA, A. B., VAN BERKEL, C. G., ROOIMANS, M. A., SCHROEDER-KURTH, T., WEGNER, R. D., GILLE, J. J., BUCHWALD, M. & ARWERT, F. 1995. Classification of Fanconi anemia patients by complementation analysis: evidence for a fifth genetic subtype. *Blood*, 86, 2156-60.
- JOENJE, H., OOSTRA, A. B., WIJKER, M., DI SUMMA, F. M., VAN BERKEL, C. G., ROOIMANS, M. A., EBELL, W., VAN WEEL, M., PRONK, J. C., BUCHWALD, M. & ARWERT, F. 1997. Evidence for at least eight Fanconi anemia genes. *Am J Hum Genet*, 61, 940-4.
- JOENJE, H. & PATEL, K. J. 2001. The emerging genetic and molecular basis of Fanconi anaemia. *Nat Rev Genet*, 2, 446-57.
- KADDAR, T. & CARREAU, M. 2012. Fanconi anemia proteins and their interacting partners: a molecular puzzle. *Anemia*, 2012, 425814.
- KALB, R., NEVELING, K., HOEHN, H., SCHNEIDER, H., LINKA, Y., BATISH, S. D., HUNT, C., BERWICK, M., CALLEN, E., SURRALLES, J., CASADO, J. A., BUEREN, J., DASI, A., SOULIER, J., GLUCKMAN, E., ZWAAN, C. M., VAN SPAENDONK, R., PALS, G., DE WINTER, J. P., JOENJE, H., GROMPE, M., AUERBACH, A. D., HANENBERG, H. & SCHINDLER, D. 2007. Hypomorphic mutations in the gene encoding a key Fanconi anemia protein, FANCD2, sustain a significant group of FA-D2 patients with severe phenotype. *Am J Hum Genet*, 80, 895-910.

- KALB, R., NEVELING, K., NANDA, I., SCHINDLER, D. & HOEHN, H. 2006. Fanconi anemia: causes and consequences of genetic instability. *Genome Dyn*, 1, 218-42.
- KAPOSI, M. 1883. Xeroderma pigmentosum. *Ann Dermatol Venereol.*, 4, 29–38.
- KAROW, J. K., CONSTANTINO, A., LI, J. L., WEST, S. C. & HICKSON, I. D. 2000. The Bloom's syndrome gene product promotes branch migration of holliday junctions. *Proc Natl Acad Sci U S A*, 97, 6504-8.
- KASHIYAMA, K., NAKAZAWA, Y., PILZ, D. T., GUO, C., SHIMADA, M., SASAKI, K., FAWCETT, H., WING, J. F., LEWIN, S. O., CARR, L., LI, T. S., YOSHIURA, K., UTANI, A., HIRANO, A., YAMASHITA, S., GREENBLATT, D., NARDO, T., STEFANINI, M., MCGIBBON, D., SARKANY, R., FASSIHI, H., TAKAHASHI, Y., NAGAYAMA, Y., MITSUTAKE, N., LEHMANN, A. R. & OGI, T. 2013. Malfunction of nuclease ERCC1-XPF results in diverse clinical manifestations and causes Cockayne syndrome, xeroderma pigmentosum, and Fanconi anemia. *Am J Hum Genet*, 92, 807-19.
- KIISKI, J. I., PELTTARI, L. M., KHAN, S., FREYSTEINSDOTTIR, E. S., REYNISDOTTIR, I., HART, S. N., SHIMELIS, H., VILSKE, S., KALLIONIEMI, A., SCHLEUTKER, J., LEMINEN, A., BUTZOW, R., BLOMQUIST, C., BARKARDOTTIR, R. B., COUCH, F. J., AITTO MAKI, K. & NEVANLINNA, H. 2014. Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. *Proc Natl Acad Sci U S A*, 111, 15172-7.
- KIM, H., CHEN, J. & YU, X. 2007a. Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science*, 316, 1202-5.
- KIM, J. B., PORRECA, G. J., SONG, L., GREENWAY, S. C., GORHAM, J. M., CHURCH, G. M., SEIDMAN, C. E. & SEIDMAN, J. G. 2007b. Polony multiplex analysis of gene expression (PMAGE) in mouse hypertrophic cardiomyopathy. *Science*, 316, 1481-4.
- KIM, J. M., PARMAR, K., HUANG, M., WEINSTOCK, D. M., RUIT, C. A., KUTOK, J. L. & D'ANDREA, A. D. 2009. Inactivation of murine Usp1 results in genomic instability and a Fanconi anemia phenotype. *Dev Cell*, 16, 314-20.
- KIM, Y., LACH, F. P., DESETTY, R., HANENBERG, H., AUERBACH, A. D. & SMOGORZEWSKA, A. 2011. Mutations of the SLX4 gene in Fanconi anemia. *Nat Genet*, 43, 142-6.
- KLEIJER, W. J., LAUGEL, V., BERNEBURG, M., NARDO, T., FAWCETT, H., GRATCHEV, A., JASPERS, N. G., SARASIN, A., STEFANINI, M. & LEHMANN, A. R. 2008. Incidence of DNA repair deficiency disorders in western Europe: Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *DNA Repair (Amst)*, 7, 744-50.
- KLEIJER, W. J., VAN DER STERRE, M. L., GARRITSEN, V. H., RAAMS, A. & JASPERS, N. G. 2006. Prenatal diagnosis of the Cockayne syndrome: survey of 15 years experience. *Prenat Diagn*, 26, 980-4.
- KLEIN DOUWEL, D., BOONEN, R. A., LONG, D. T., SZYPOWSKA, A. A., RASCHLE, M., WALTER, J. C. & KNIPSCHER, P. 2014. XPF-ERCC1 acts in Unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Mol Cell*, 54, 460-71.
- KNIES, K., SCHUSTER, B., AMEZIANE, N., ROOIMANS, M., BETTECKEN, T., DE WINTER, J. & SCHINDLER, D. 2012. Genotyping of fanconi anemia patients by whole exome sequencing: advantages and challenges. *PLoS One*, 7, e52648.
- KNIPSCHER, P., RASCHLE, M., SMOGORZEWSKA, A., ENOIU, M., HO, T. V., SCHARER, O. D., ELLEDGE, S. J. & WALTER, J. C. 2009. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science*, 326, 1698-701.
- KOHLHASE, S., BOGDANOVA, N. V., SCHURMANN, P., BERMISHEVA, M., KHUSNUTDINOVA, E., ANTONENKOVA, N., PARK-SIMON, T. W., HILLEMANN, P., MEYER, A., CHRISTIANSEN, H., SCHINDLER, D. & DORK, T. 2014. Mutation analysis of the ERCC4/FANCP gene in hereditary breast cancer. *PLoS One*, 9, e85334.

- KORLACH, J., BJORNSON, K. P., CHAUDHURI, B. P., CICERO, R. L., FLUSBERG, B. A., GRAY, J. J., HOLDEN, D., SAXENA, R., WEGENER, J. & TURNER, S. W. 2010. Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol*, 472, 431-55.
- KRATZ, K., SCHOPF, B., KADEN, S., SENDOEL, A., EBERHARD, R., LADEMANN, C., CANNAVO, E., SARTORI, A. A., HENGARTNER, M. O. & JIRICNY, J. 2010. Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. *Cell*, 142, 77-88.
- KU, C. S., NAIDOO, N. & PAWITAN, Y. 2011. Revisiting Mendelian disorders through exome sequencing. *Hum Genet*, 129, 351-70.
- KUBBIES, M., SCHINDLER, D., HOEHN, H., SCHINZEL, A. & RABINOVITCH, P. S. 1985. Endogenous blockage and delay of the chromosome cycle despite normal recruitment and growth phase explain poor proliferation and frequent edomitosis in Fanconi anemia cells. *Am J Hum Genet*, 37, 1022-30.
- KUMAR, P., HENIKOFF, S. & NG, P. C. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*, 4, 1073-81.
- KURAOKA, I., KOBERTZ, W. R., ARIZA, R. R., BIGGERSTAFF, M., ESSIGMANN, J. M. & WOOD, R. D. 2000. Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease. *J Biol Chem*, 275, 26632-6.
- KUTLER, D. I. & AUERBACH, A. D. 2004. Fanconi anemia in Ashkenazi Jews. *Fam Cancer*, 3, 241-8.
- KUTLER, D. I., SINGH, B., SATAGOPAN, J., BATISH, S. D., BERWICK, M., GIAMPIETRO, P. F., HANENBERG, H. & AUERBACH, A. D. 2003. A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood*, 101, 1249-56.
- KWEE, M. L., POLL, E. H., VAN DE KAMP, J. J., DE KONING, H., ERIKSSON, A. W. & JOENJE, H. 1983. Unusual response to bifunctional alkylating agents in a case of Fanconi anaemia. *Hum Genet*, 64, 384-7.
- LANDWEHR, R., BOGDANOVA, N. V., ANTONENKOVA, N., MEYER, A., BREMER, M., PARK-SIMON, T. W., HILLEMANN, P., KARSTENS, J. H., SCHINDLER, D. & DORK, T. 2011. Mutation analysis of the SLX4/FANCP gene in hereditary breast cancer. *Breast Cancer Res Treat*, 130, 1021-8.
- LANGEVIN, F., CROSSAN, G. P., ROSADO, I. V., ARENDS, M. J. & PATEL, K. J. 2011. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature*, 475, 53-8.
- LANS, H., MARTEIJN, J. A. & VERMEULEN, W. 2012. ATP-dependent chromatin remodeling in the DNA-damage response. *Epigenetics Chromatin*, 5, 4.
- LAVIN, M. F. 2007. ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks. *Oncogene*, 26, 7749-58.
- LAVIN, M. F. & KOZLOV, S. 2007. ATM activation and DNA damage response. *Cell Cycle*, 6, 931-42.
- LEAMON, J. H., LEE, W. L., TARTARO, K. R., LANZA, J. R., SARKIS, G. J., DEWINTER, A. D., BERKA, J., WEINER, M., ROTHBERG, J. M. & LOHMAN, K. L. 2003. A massively parallel PicoTiterPlate based platform for discrete picoliter-scale polymerase chain reactions. *Electrophoresis*, 24, 3769-77.
- LEE, J. H. & PAULL, T. T. 2007. Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene*, 26, 7741-8.
- LEHMANN, A. R. 2003. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie*, 85, 1101-11.
- LEHMANN, A. R., BOOTSMA, D., CLARKSON, S. G., CLEAVER, J. E., MCALPINE, P. J., TANAKA, K., THOMPSON, L. H. & WOOD, R. D. 1994. Nomenclature of human DNA repair genes. *Mutat Res*, 315, 41-2.
- LEHMANN, A. R., MCGIBBON, D. & STEFANINI, M. 2011. Xeroderma pigmentosum. *Orphanet J Rare Dis*, 6, 70.

- LEVITUS, M., ROOIMANS, M. A., STELTENPOOL, J., COOL, N. F., OOSTRA, A. B., MATHEW, C. G., HOATLIN, M. E., WAISFISZ, Q., ARWERT, F., DE WINTER, J. P. & JOENJE, H. 2004. Heterogeneity in Fanconi anemia: evidence for 2 new genetic subtypes. *Blood*, 103, 2498-503.
- LEVITUS, M., WAISFISZ, Q., GODTHELP, B. C., DE VRIES, Y., HUSSAIN, S., WIEGANT, W. W., ELGHALBZOURI-MAGHRANI, E., STELTENPOOL, J., ROOIMANS, M. A., PALS, G., ARWERT, F., MATHEW, C. G., ZDZIENICKA, M. Z., HIOM, K., DE WINTER, J. P. & JOENJE, H. 2005. The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. *Nat Genet*, 37, 934-5.
- LEVRAN, O., ATTWOOLL, C., HENRY, R. T., MILTON, K. L., NEVELING, K., RIO, P., BATISH, S. D., KALB, R., VELLEUER, E., BARRAL, S., OTT, J., PETRINI, J., SCHINDLER, D., HANENBERG, H. & AUERBACH, A. D. 2005a. The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. *Nat Genet*, 37, 931-3.
- LEVRAN, O., DIOTTI, R., PUJARA, K., BATISH, S. D., HANENBERG, H. & AUERBACH, A. D. 2005b. Spectrum of sequence variations in the FANCA gene: an International Fanconi Anemia Registry (IFAR) study. *Hum Mutat*, 25, 142-9.
- LEVY-LAHAD, E. 2010. Fanconi anemia and breast cancer susceptibility meet again. *Nat Genet*, 42, 368-9.
- LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25, 1754-60.
- LI, H., HANDSAKER, B., WYSOKER, A., FENNEL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G. & DURBIN, R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25, 2078-9.
- LIM, E. T., WURTZ, P., HAVULINNA, A. S., PALTA, P., TUKIAINEN, T., REHNSTROM, K., ESKO, T., MAGI, R., INOUE, M., LAPPALAINEN, T., CHAN, Y., SALEM, R. M., LEK, M., FLANNICK, J., SIM, X., MANNING, A., LADENVALL, C., BUMPSTEAD, S., HAMALAINEN, E., AALTO, K., MAKSIMOW, M., SALMI, M., BLANKENBERG, S., ARDISSINO, D., SHAH, S., HORNE, B., MCPHERSON, R., HOVINGH, G. K., REILLY, M. P., WATKINS, H., GOEL, A., FARRALL, M., GIRELLI, D., REINER, A. P., STITZEL, N. O., KATHIRESAN, S., GABRIEL, S., BARRETT, J. C., LEHTIMAKI, T., LAAKSO, M., GROOP, L., KAPRIO, J., PEROLA, M., MCCARTHY, M. I., BOEHNKE, M., ALTSHULER, D. M., LINDGREN, C. M., HIRSCHHORN, J. N., METSPALU, A., FREIMER, N. B., ZELLER, T., JALKANEN, S., KOSKINEN, S., RAITAKARI, O., DURBIN, R., MACARTHUR, D. G., SALOMAA, V., RIPATTI, S., DALY, M. J. & PALOTIE, A. 2014. Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet*, 10, e1004494.
- LIMSIRICHAIKUL, S., NIIMI, A., FAWCETT, H., LEHMANN, A., YAMASHITA, S. & OGI, T. 2009. A rapid non-radioactive technique for measurement of repair synthesis in primary human fibroblasts by incorporation of ethynyl deoxyuridine (EdU). *Nucleic Acids Res*, 37, e31.
- LITMAN, R., PENG, M., JIN, Z., ZHANG, F., ZHANG, J., POWELL, S., ANDREASSEN, P. R. & CANTOR, S. B. 2005. BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCF. *Cancer Cell*, 8, 255-65.
- LIU, T., GHOSAL, G., YUAN, J., CHEN, J. & HUANG, J. 2010. FAN1 acts with FANCF-FANCD2 to promote DNA interstrand cross-link repair. *Science*, 329, 693-6.
- LIU, T. X., HOWLETT, N. G., DENG, M., LANGENAU, D. M., HSU, K., RHODES, J., KANKI, J. P., D'ANDREA, A. D. & LOOK, A. T. 2003. Knockdown of zebrafish Fancd2 causes developmental abnormalities via p53-dependent apoptosis. *Dev Cell*, 5, 903-14.
- LIVINGSTON, D. M. 2009. Cancer. Complicated supercomplexes. *Science*, 324, 602-3.
- LO TEN FOE, J. R., KWEE, M. L., ROOIMANS, M. A., OOSTRA, A. B., VEERMAN, A. J., VAN WEEL, M., PAULI, R. M., SHAHIDI, N. T., DOKAL, I., ROBERTS, I., ALTAY, C., GLUCKMAN, E., GIBSON, R.

- A., MATHEW, C. G., ARWERT, F. & JOENJE, H. 1997. Somatic mosaicism in Fanconi anemia: molecular basis and clinical significance. *Eur J Hum Genet*, 5, 137-48.
- LO TEN FOE, J. R., ROOIMANS, M. A., BOSNOYAN-COLLINS, L., ALON, N., WIJKER, M., PARKER, L., LIGHTFOOT, J., CARREAU, M., CALLEN, D. F., SAVOIA, A., CHENG, N. C., VAN BERKEL, C. G., STRUNK, M. H., GILLE, J. J., PALS, G., KRUYT, F. A., PRONK, J. C., ARWERT, F., BUCHWALD, M. & JOENJE, H. 1996. Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA. *Nat Genet*, 14, 320-3.
- LOBER, G. & KITTLER, L. 1977. Selected topics in photochemistry of nucleic acids. Recent results and perspectives. *Photochem Photobiol*, 25, 215-33.
- LONGERICH, S., SAN FILIPPO, J., LIU, D. & SUNG, P. 2009. FANCI binds branched DNA and is monoubiquitinated by UBE2T-FANCL. *J Biol Chem*, 284, 23182-6.
- LOVEDAY, C., TURNBULL, C., RUARK, E., XICOLA, R. M., RAMSAY, E., HUGHES, D., WARREN-PERRY, M., SNAPE, K., ECCLES, D., EVANS, D. G., GORE, M., RENWICK, A., SEAL, S., ANTONIOU, A. C. & RAHMAN, N. 2012. Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nat Genet*, 44, 475-6; author reply 476.
- MACKAY, C., DECLAIS, A. C., LUNDIN, C., AGOSTINHO, A., DEANS, A. J., MACARTNEY, T. J., HOFMANN, K., GARTNER, A., WEST, S. C., HELLEDAY, T., LILLEY, D. M. & ROUSE, J. 2010. Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell*, 142, 65-76.
- MACHIDA, Y. J., MACHIDA, Y., CHEN, Y., GURTAN, A. M., KUPFER, G. M., D'ANDREA, A. D. & DUTTA, A. 2006. UBE2T is the E2 in the Fanconi anemia pathway and undergoes negative autoregulation. *Mol Cell*, 23, 589-96.
- MAGDALENA, N., PILONETTO, D. V., BITENCOURT, M. A., PEREIRA, N. F., RIBEIRO, R. C., JENG, M. & PASQUINI, R. 2005. Frequency of Fanconi anemia in Brazil and efficacy of screening for the FANCA 3788-3790del mutation. *Braz J Med Biol Res*, 38, 669-73.
- MANANDHAR, M., BOULWARE, K. S. & WOOD, R. D. 2015. The ERCC1 and ERCC4 (XPF) genes and gene products. *Gene*, 569, 153-61.
- MASSINGHAM, T. & GOLDMAN, N. 2012. Error-correcting properties of the SOLiD Exact Call Chemistry. *BMC Bioinformatics*, 13, 145.
- MASSON, J. Y., TARSOUNAS, M. C., STASIAK, A. Z., STASIAK, A., SHAH, R., MCILWRAITH, M. J., BENSON, F. E. & WEST, S. C. 2001. Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev*, 15, 3296-307.
- MATSUOKA, S., BALLIF, B. A., SMOGORZEWSKA, A., MCDONALD, E. R., 3RD, HUROV, K. E., LUO, J., BAKALARSKI, C. E., ZHAO, Z., SOLIMINI, N., LERENTHAL, Y., SHILOH, Y., GYGI, S. P. & ELLEDGE, S. J. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, 316, 1160-6.
- MCCAULEY, J., MASAND, N., MCGOWAN, R., RAJAGOPALAN, S., HUNTER, A., MICHAUD, J. L., GIBSON, K., ROBERTSON, J., VAZ, F., ABBS, S. & HOLDEN, S. T. 2011. X-linked VACTERL with hydrocephalus syndrome: further delineation of the phenotype caused by FANCB mutations. *Am J Med Genet A*, 155A, 2370-80.
- MCILWRAITH, M. J., VAISMAN, A., LIU, Y., FANNING, E., WOODGATE, R. & WEST, S. C. 2005. Human DNA polymerase eta promotes DNA synthesis from strand invasion intermediates of homologous recombination. *Mol Cell*, 20, 783-92.
- MCILWRAITH, M. J., VAN DYCK, E., MASSON, J. Y., STASIAK, A. Z., STASIAK, A. & WEST, S. C. 2000. Reconstitution of the strand invasion step of double-strand break repair using human Rad51 Rad52 and RPA proteins. *J Mol Biol*, 304, 151-64.
- MCILWRAITH, M. J. & WEST, S. C. 2008. DNA repair synthesis facilitates RAD52-mediated second-end capture during DSB repair. *Mol Cell*, 29, 510-6.

- MCKENNA, A., HANNA, M., BANKS, E., SIVACHENKO, A., CIBULSKIS, K., KERNYTSKY, A., GARIMELLA, K., ALTSHULER, D., GABRIEL, S., DALY, M. & DEPRISTO, M. A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*, 20, 1297-303.
- MCWHIR, J., SELFRIDGE, J., HARRISON, D. J., SQUIRES, S. & MELTON, D. W. 1993. Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat Genet*, 5, 217-24.
- MEETEI, A. R., DE WINTER, J. P., MEDHURST, A. L., WALLISCH, M., WAISFISZ, Q., VAN DE VRUGT, H. J., OOSTRA, A. B., YAN, Z., LING, C., BISHOP, C. E., HOATLIN, M. E., JOENJE, H. & WANG, W. 2003. A novel ubiquitin ligase is deficient in Fanconi anemia. *Nat Genet*, 35, 165-70.
- MEETEI, A. R., LEVITUS, M., XUE, Y., MEDHURST, A. L., ZWAAN, M., LING, C., ROOIMANS, M. A., BIER, P., HOATLIN, M., PALS, G., DE WINTER, J. P., WANG, W. & JOENJE, H. 2004. X-linked inheritance of Fanconi anemia complementation group B. *Nat Genet*, 36, 1219-24.
- MEETEI, A. R., MEDHURST, A. L., LING, C., XUE, Y., SINGH, T. R., BIER, P., STELTENPOOL, J., STONE, S., DOKAL, I., MATHEW, C. G., HOATLIN, M., JOENJE, H., DE WINTER, J. P. & WANG, W. 2005. A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet*, 37, 958-63.
- MEINDL, A., HELLEBRAND, H., WIEK, C., ERVEN, V., WAPPENSCHMIDT, B., NIEDERACHER, D., FREUND, M., LICHTNER, P., HARTMANN, L., SCHAAL, H., RAMSER, J., HONISCH, E., KUBISCH, C., WICHMANN, H. E., KAST, K., DEISSLER, H., ENGEL, C., MULLER-MYHSOK, B., NEVELING, K., KIECHLE, M., MATHEW, C. G., SCHINDLER, D., SCHMUTZLER, R. K. & HANENBERG, H. 2010. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat Genet*, 42, 410-4.
- METZKER, M. L. 2010. Sequencing technologies - the next generation. *Nat Rev Genet*, 11, 31-46.
- MEYER, S., TISCHKOWITZ, M., CHANDLER, K., GILLESPIE, A., BIRCH, J. M. & EVANS, D. G. 2014. Fanconi anaemia, BRCA2 mutations and childhood cancer: a developmental perspective from clinical and epidemiological observations with implications for genetic counselling. *J Med Genet*, 51, 71-5.
- MEYNERT, A. M., BICKNELL, L. S., HURLES, M. E., JACKSON, A. P. & TAYLOR, M. S. 2013. Quantifying single nucleotide variant detection sensitivity in exome sequencing. *BMC Bioinformatics*, 14, 195.
- MI, J., QIAO, F., WILSON, J. B., HIGH, A. A., SCHROEDER, M. J., STUKENBERG, P. T., MOSS, A., SHABANOWITZ, J., HUNT, D. F., JONES, N. J. & KUPFER, G. M. 2004. FANCG is phosphorylated at serines 383 and 387 during mitosis. *Mol Cell Biol*, 24, 8576-85.
- MILNE, R. L., RIBAS, G., GONZALEZ-NEIRA, A., FAGERHOLM, R., SALAS, A., GONZALEZ, E., DOPAZO, J., NEVANLINNA, H., ROBLEDO, M. & BENITEZ, J. 2006. ERCC4 associated with breast cancer risk: a two-stage case-control study using high-throughput genotyping. *Cancer Res*, 66, 9420-7.
- MILOS, P. M. 2009. Emergence of single-molecule sequencing and potential for molecular diagnostic applications. *Expert Rev Mol Diagn*, 9, 659-66.
- MIMITOU, E. P. & SYMINGTON, L. S. 2009. Nucleases and helicases take center stage in homologous recombination. *Trends Biochem Sci*, 34, 264-72.
- MITRA, R. D., SHENDURE, J., OLEJNIK, J., EDYTA KRZYMANSKA, O. & CHURCH, G. M. 2003. Fluorescent in situ sequencing on polymerase colonies. *Anal Biochem*, 320, 55-65.
- MOGGS, J. G., YAREMA, K. J., ESSIGMANN, J. M. & WOOD, R. D. 1996. Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *J Biol Chem*, 271, 7177-86.

- MORGAN, N. V., TIPPING, A. J., JOENJE, H. & MATHEW, C. G. 1999. High frequency of large intragenic deletions in the Fanconi anemia group A gene. *Am J Hum Genet*, 65, 1330-41.
- MOSEDALE, G., NIEDZWIEDZ, W., ALPI, A., PERRINA, F., PEREIRA-LEAL, J. B., JOHNSON, M., LANGEVIN, F., PACE, P. & PATEL, K. J. 2005. The vertebrate Hef ortholog is a component of the Fanconi anemia tumor-suppressor pathway. *Nat Struct Mol Biol*, 12, 763-71.
- MU, D., HSU, D. S. & SANCAR, A. 1996. Reaction mechanism of human DNA repair excision nuclease. *J Biol Chem*, 271, 8285-94.
- MUNOZ, I. M., HAIN, K., DECLAIS, A. C., GARDINER, M., TOH, G. W., SANCHEZ-PULIDO, L., HEUCKMANN, J. M., TOTH, R., MACARTNEY, T., EPPINK, B., KANAAR, R., PONTING, C. P., LILLEY, D. M. & ROUSE, J. 2009. Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Mol Cell*, 35, 116-27.
- NABHAN, S. K., BITENCOURT, M. A., DUVAL, M., ABECASIS, M., DUFOUR, C., BOUDJEDIR, K., ROCHA, V., SOCIE, G., PASSWEG, J., GOI, K., SANDERS, J., SNOWDEN, J., YABE, H., PASQUINI, R., GLUCKMAN, E. & APLASTIC ANAEMIA WORKING PARTY, E. 2010. Fertility recovery and pregnancy after allogeneic hematopoietic stem cell transplantation in Fanconi anemia patients. *Haematologica*, 95, 1783-7.
- NAKANO, M., KOMATSU, J., MATSUURA, S., TAKASHIMA, K., KATSURA, S. & MIZUNO, A. 2003. Single-molecule PCR using water-in-oil emulsion. *J Biotechnol*, 102, 117-24.
- NASSIF, N., PENNEY, J., PAL, S., ENGELS, W. R. & GLOOR, G. B. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol*, 14, 1613-25.
- NEW, J. H., SUGIYAMA, T., ZAITSEVA, E. & KOWALCZYKOWSKI, S. C. 1998. Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature*, 391, 407-10.
- NG, S. B., TURNER, E. H., ROBERTSON, P. D., FLYGARE, S. D., BIGHAM, A. W., LEE, C., SHAFFER, T., WONG, M., BHATTACHARJEE, A., EICHLER, E. E., BAMSHAD, M., NICKERSON, D. A. & SHENDURE, J. 2009. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, 461, 272-6.
- NICCHIA, E., GRECO, C., DE ROCCO, D., PECILE, V., D'EUSTACCHIO, A., CAPPELLI, E., CORTI, P., MARRA, N., RAMENGI, U., PILLON, M., FARRUGGIA, P., DUFOUR, C., PALLAVICINI, A., TORELLI, L. & SAVOIA, A. 2015. Identification of point mutations and large intragenic deletions in Fanconi anemia using next-generation sequencing technology. *Molecular Genetics & Genomic Medicine*, n/a-n/a.
- NIEDERNHOFER, L. J., GARINIS, G. A., RAAMS, A., LALAI, A. S., ROBINSON, A. R., APPELDOORN, E., ODIJK, H., OOSTENDORP, R., AHMAD, A., VAN LEEUWEN, W., THEIL, A. F., VERMEULEN, W., VAN DER HORST, G. T., MEINECKE, P., KLEIJER, W. J., VIJG, J., JASPERS, N. G. & HOEIJMAKERS, J. H. 2006. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature*, 444, 1038-43.
- NIEDZWIEDZ, W., MOSEDALE, G., JOHNSON, M., ONG, C. Y., PACE, P. & PATEL, K. J. 2004. The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Mol Cell*, 15, 607-20.
- NIJMAN, S. M., HUANG, T. T., DIRAC, A. M., BRUMMELKAMP, T. R., KERKHOVEN, R. M., D'ANDREA, A. D. & BERNARDS, R. 2005. The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. *Mol Cell*, 17, 331-9.
- NIMONKAR, A. V., OZSOY, A. Z., GENSCHER, J., MODRICH, P. & KOWALCZYKOWSKI, S. C. 2008. Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. *Proc Natl Acad Sci U S A*, 105, 16906-11.
- NOJIMA, K., HOCHEGGER, H., SABERI, A., FUKUSHIMA, T., KIKUCHI, K., YOSHIMURA, M., ORELLI, B. J., BISHOP, D. K., HIRANO, S., OHZEKI, M., ISHIAI, M., YAMAMOTO, K., TAKATA, M., ARAKAWA,

- H., BUERSTEDDE, J. M., YAMAZOE, M., KAWAMOTO, T., ARAKI, K., TAKAHASHI, J. A., HASHIMOTO, N., TAKEDA, S. & SONODA, E. 2005. Multiple repair pathways mediate tolerance to chemotherapeutic cross-linking agents in vertebrate cells. *Cancer Res*, 65, 11704-11.
- NOUSPIKEL, T., LALLE, P., LEADON, S. A., COOPER, P. K. & CLARKSON, S. G. 1997. A common mutational pattern in Cockayne syndrome patients from xeroderma pigmentosum group G: implications for a second XPG function. *Proc Natl Acad Sci U S A*, 94, 3116-21.
- O'BRIEN, P. J., SIRAKI, A. G. & SHANGARI, N. 2005. Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. *Crit Rev Toxicol*, 35, 609-62.
- OBERBECK, N., LANGEVIN, F., KING, G., DE WIND, N., CROSSAN, G. P. & PATEL, K. J. 2014. Maternal aldehyde elimination during pregnancy preserves the fetal genome. *Mol Cell*, 55, 807-17.
- OESTERGAARD, V. H., LANGEVIN, F., KUIKEN, H. J., PACE, P., NIEDZWIEDZ, W., SIMPSON, L. J., OHZEKI, M., TAKATA, M., SALE, J. E. & PATEL, K. J. 2007. Deubiquitination of FANCD2 is required for DNA crosslink repair. *Mol Cell*, 28, 798-809.
- OLIVER, A. W., SWIFT, S., LORD, C. J., ASHWORTH, A. & PEARL, L. H. 2009. Structural basis for recruitment of BRCA2 by PALB2. *EMBO Rep*, 10, 990-6.
- ORELLI, B., MCCLENDON, T. B., TSODIKOV, O. V., ELLENBERGER, T., NIEDERNHOFER, L. J. & SCHARER, O. D. 2010. The XPA-binding domain of ERCC1 is required for nucleotide excision repair but not other DNA repair pathways. *J Biol Chem*, 285, 3705-12.
- OSORIO, A., BOGLIOLO, M., FERNÁNDEZ, V., BARROSO, A., DE LA HOYA, M., CALDÉS, T., LASA, A., CAJAL, T. R. Y., SANTAMARIÑA, M., VEGA, A., QUILES, F., LÁZARO, C., DÍEZ, O., FERNÁNDEZ, D., GONZÁLEZ-SARMIENTO, R., DURÁN, M., PIQUERAS, J. F., MARÍN, M., PUJOL, R., SURRALLÉS, J. & BENÍTEZ, J. 2013. Evaluation of Rare Variants in the New Fanconi Anemia Gene ERCC4 (FANCF) as Familial Breast/Ovarian Cancer Susceptibility Alleles. *Hum Mutat*, n/a-n/a.
- OSORIO, A., ENDT, D., FERNANDEZ, F., EIRICH, K., DE LA HOYA, M., SCHMUTZLER, R., CALDES, T., MEINDL, A., SCHINDLER, D. & BENITEZ, J. 2012. Predominance of pathogenic missense variants in the RAD51C gene occurring in breast and ovarian cancer families. *Hum Mol Genet*, 21, 2889-98.
- PANIER, S. & DUROCHER, D. 2009. Regulatory ubiquitylation in response to DNA double-strand breaks. *DNA Repair (Amst)*, 8, 436-443.
- PARMAR, K., D'ANDREA, A. & NIEDERNHOFER, L. J. 2009. Mouse models of Fanconi anemia. *Mutat Res*, 668, 133-40.
- PARMAR, K. & D'ANDREA, A. D. 2012. Stressed out: endogenous aldehydes damage hematopoietic stem cells. *Cell Stem Cell*, 11, 583-4.
- PATTENGALE, P. K., SMITH, R. W. & GERBER, P. 1973. Selective transformation of B lymphocytes by E.B. virus. *Lancet*, 2, 93-4.
- PELLETIER, J. & SONENBERG, N. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, 334, 320-5.
- PETERLONGO, P., CATUCCI, I., COLOMBO, M., CALECA, L., MUCAKI, E., BOGLIOLO, M., MARIN, M., DAMIOLA, F., BERNARD, L., PENSOTTI, V., VOLORIO, S., DALL'OLIO, V., MEINDL, A., BARTRAM, C., SUTTER, C., SUROWY, H., SORNIN, V., DONDON, M. G., EON-MARCHAIS, S., STOPPA-LYONNET, D., ANDRIEU, N., SINILNIKOVA, O. M., GENESIS, MITCHELL, G., JAMES, P. A., THOMPSON, E., KCONFAB, SWE, B., MARCHETTI, M., VERZEROLI, C., TARTARI, C., CAPONE, G., PUTIGNANO, A. L., GENUARDI, M., MEDICI, V., MARCHI, I., FEDERICO, M., TOGNAZZO, S., MATRICARDI, L., AGATA, S., DOLCETTI, R., DELLA PUPPA, L., CINI, G., GISMONDI, V., VIASSOLO, V., PERFUMO, C., MENCARELLI, M. A., BALDASSARRI, M., PEISSEL, B., ROVERSI, G., SILVESTRI, V., RIZZOLO, P., SPINA, F., VIVANET, C., TIBILETTI, M. G., CALIGO,

- M. A., GAMBINO, G., TOMMASI, S., PILATO, B., TONDINI, C., CORNA, C., BONANNI, B., BARILE, M., OSORIO, A., BENITEZ, J., BALESTRINO, L., OTTINI, L., MANOUKIAN, S., PIEROTTI, M. A., RENIERI, A., VARESCO, L., COUCH, F. J., WANG, X., DEVILEE, P., HILBERS, F. S., VAN ASPEREN, C. J., VIEL, A., MONTAGNA, M., CORTESI, L., DIEZ, O., BALMANA, J., HAUKE, J., SCHMUTZLER, R. K., PAPI, L., PUJANA, M. A., LAZARO, C., FALANGA, A., OFFIT, K., VIJAI, J., CAMPBELL, I., BURWINKEL, B., KVIST, A., EHRENCRONA, H., MAZOYER, S., PIZZAMIGLIO, S., VERDERIO, P., SURRALLES, J., ROGAN, P. K. & RADICE, P. 2015. FANCM c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity, and is a familial breast cancer risk factor. *Hum Mol Genet*.
- PETRUCCELLI, N., DALY, M. B. & FELDMAN, G. L. 2010. Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. *Genet Med*, 12, 245-59.
- PINTO, F. O., LEBLANC, T., CHAMOUSSET, D., LE ROUX, G., BRETHON, B., CASSINAT, B., LARGHERO, J., DE VILLARTAY, J. P., STOPPA-LYONNET, D., BARUCHEL, A., SOCIE, G., GLUCKMAN, E. & SOULIER, J. 2009. Diagnosis of Fanconi anemia in patients with bone marrow failure. *Haematologica*, 94, 487-95.
- PRASHER, J. M., LALAI, A. S., HEIJMANS-ANTONISSEN, C., PLOEMACHER, R. E., HOEIJMAKERS, J. H., TOUW, I. P. & NIEDERNHOFER, L. J. 2005. Reduced hematopoietic reserves in DNA interstrand crosslink repair-deficient *Ercc1*^{-/-} mice. *EMBO J*, 24, 861-71.
- QIAO, F., MI, J., WILSON, J. B., ZHI, G., BUCHEIMER, N. R., JONES, N. J. & KUPFER, G. M. 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, 46035-45.
- QUENTIN, S., CUCCUINI, W., CECCALDI, R., NIBOUREL, O., PONDARRE, C., PAGES, M. P., VASQUEZ, N., DUBOIS D'ENGLISHEN, C., LARGHERO, J., PEFFAULT DE LATOUR, R., ROCHA, V., DALLE, J. H., SCHNEIDER, P., MICHALLET, M., MICHEL, G., BARUCHEL, A., SIGAUX, F., GLUCKMAN, E., LEBLANC, T., STOPPA-LYONNET, D., PREUDHOMME, C., SOCIE, G. & SOULIER, J. 2011. Myelodysplasia and leukemia of Fanconi anemia are associated with a specific pattern of genomic abnormalities that includes cryptic RUNX1/AML1 lesions. *Blood*, 117, e161-70.
- RAHN, J. J., ADAIR, G. M. & NAIRN, R. S. 2010. Multiple roles of ERCC1-XPF in mammalian interstrand crosslink repair. *Environ Mol Mutagen*, 51, 567-81.
- RASCHLE, M., KNIPSCHER, P., ENOIU, M., ANGELOV, T., SUN, J., GRIFFITH, J. D., ELLENBERGER, T. E., SCHARER, O. D. & WALTER, J. C. 2008. Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell*, 134, 969-80.
- REID, S., SCHINDLER, D., HANENBERG, H., BARKER, K., HANKS, S., KALB, R., NEVELING, K., KELLY, P., SEAL, S., FREUND, M., WURM, M., BATISH, S. D., LACH, F. P., YETGIN, S., NEITZEL, H., ARIFFIN, H., TISCHKOWITZ, M., MATHEW, C. G., AUERBACH, A. D. & RAHMAN, N. 2007. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat Genet*, 39, 162-4.
- RICKMAN, K. A., LACH, F. P., ABHYANKAR, A., DONOVAN, F. X., SANBORN, E. M., KENNEDY, J. A., SOUGNEZ, C., GABRIEL, S. B., ELEMENTO, O., CHANDRASEKHARAPPA, S. C., SCHINDLER, D., AUERBACH, A. D. & SMOGORZEWSKA, A. 2015. Deficiency of UBE2T, the E2 Ubiquitin Ligase Necessary for FANCD2 and FANCI Ubiquitination, Causes FA-T Subtype of Fanconi Anemia. *Cell Rep*, 12, 35-41.
- RIO, P., BANOS, R., LOMBARDO, A., QUINTANA-BUSTAMANTE, O., ALVAREZ, L., GARATE, Z., GENOVESE, P., ALMARZA, E., VALERI, A., DIEZ, B., NAVARRO, S., TORRES, Y., TRUJILLO, J. P., MURILLAS, R., SEGOVIA, J. C., SAMPER, E., SURRALLES, J., GREGORY, P. D., HOLMES, M. C., NALDINI, L. & BUEREN, J. A. 2014. Targeted gene therapy and cell reprogramming in Fanconi anemia. *EMBO Mol Med*, 6, 835-48.

- RIOS, D., MCLAREN, W. M., CHEN, Y., BIRNEY, E., STABENAU, A., FLICEK, P. & CUNNINGHAM, F. 2010. A database and API for variation, dense genotyping and resequencing data. *BMC Bioinformatics*, 11, 238.
- RISCHEWSKI, J. & SCHNEPPENHEIM, R. 2001. Screening strategies for a highly polymorphic gene: DHPLC analysis of the Fanconi anemia group A gene. *J Biochem Biophys Methods*, 47, 53-64.
- ROBINSON, J. T., THORVALDSDOTTIR, H., WINCKLER, W., GUTTMAN, M., LANDER, E. S., GETZ, G. & MESIROV, J. P. 2011. Integrative genomics viewer. *Nat Biotechnol*, 29, 24-6.
- ROMERO, A., PEREZ-SEGURA, P., TOSAR, A., GARCIA-SAENZ, J. A., DIAZ-RUBIO, E., CALDES, T. & DE LA HOYA, M. 2011. A HRM-based screening method detects RAD51C germ-line deleterious mutations in Spanish breast and ovarian cancer families. *Breast Cancer Res Treat*, 129, 939-46.
- ROSADO, I. V., LANGEVIN, F., CROSSAN, G. P., TAKATA, M. & PATEL, K. J. 2011. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nat Struct Mol Biol*, 18, 1432-4.
- ROSENBERG, P. S., ALTER, B. P. & EBELL, W. 2008. Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry. *Haematologica*, 93, 511-7.
- ROSENBERG, P. S., GREENE, M. H. & ALTER, B. P. 2003. Cancer incidence in persons with Fanconi anemia. *Blood*, 101, 822-6.
- ROSENBERG, P. S., HUANG, Y. & ALTER, B. P. 2004. Individualized risks of first adverse events in patients with Fanconi anemia. *Blood*, 104, 350-5.
- ROTHBERG, J. M., HINZ, W., REARICK, T. M., SCHULTZ, J., MILESKI, W., DAVEY, M., LEAMON, J. H., JOHNSON, K., MILGREW, M. J., EDWARDS, M., HOON, J., SIMONS, J. F., MARRAN, D., MYERS, J. W., DAVIDSON, J. F., BRANTING, A., NOBILE, J. R., PUC, B. P., LIGHT, D., CLARK, T. A., HUBER, M., BRANCIFORTE, J. T., STONER, I. B., CAWLEY, S. E., LYONS, M., FU, Y., HOMER, N., SEDOVA, M., MIAO, X., REED, B., SABINA, J., FEIERSTEIN, E., SCHORN, M., ALANJARY, M., DIMALANTA, E., DRESSMAN, D., KASINSKAS, R., SOKOLSKY, T., FIDANZA, J. A., NAMSARAEV, E., MCKERNAN, K. J., WILLIAMS, A., ROTH, G. T. & BUSTILLO, J. 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, 475, 348-52.
- ROTHFUSS, A. & GROMPE, M. 2004. Repair kinetics of genomic interstrand DNA cross-links: evidence for DNA double-strand break-dependent activation of the Fanconi anemia/BRCA pathway. *Mol Cell Biol*, 24, 123-34.
- ROY, R., CHUN, J. & POWELL, S. N. 2012. BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat Rev Cancer*, 12, 68-78.
- SAITO, T. T., YOUNG, J. L., BOULTON, S. J. & COLAIACOVO, M. P. 2009. Caenorhabditis elegans HIM-18/SLX-4 interacts with SLX-1 and XPF-1 and maintains genomic integrity in the germline by processing recombination intermediates. *PLoS Genet*, 5, e1000735.
- SARTORI, A. A., LUKAS, C., COATES, J., MISTRICK, M., FU, S., BARTEK, J., BAER, R., LUKAS, J. & JACKSON, S. P. 2007. Human CtIP promotes DNA end resection. *Nature*, 450, 509-14.
- SASAKI, M. S. 1975. Is Fanconi's anaemia defective in a process essential to the repair of DNA cross links? *Nature*, 257, 501-3.
- SAVINO, M., BORRIELLO, A., D'APOLITO, M., CRISCUOLO, M., DEL VECCHIO, M., BIANCO, A. M., DI PERNA, M., CALZONE, R., NOBILI, B., ZATTERALE, A., ZELANTE, L., JOENJE, H., DELLA RAGIONE, F. & SAVOIA, A. 2003. Spectrum of FANCA mutations in Italian Fanconi anemia patients: identification of six novel alleles and phenotypic characterization of the S858R variant. *Hum Mutat*, 22, 338-9.
- SAVINO, M., IANZANO, L., STRIPPOLI, P., RAMENGI, U., ARSLANIAN, A., BAGNARA, G. P., JOENJE, H., ZELANTE, L. & SAVOIA, A. 1997. Mutations of the Fanconi anemia group A gene (FAA) in Italian patients. *Am J Hum Genet*, 61, 1246-53.

- SAVOIA, A., ZATTERALE, A., DEL PRINCIPE, D. & JOENJE, H. 1996. Fanconi anaemia in Italy: high prevalence of complementation group A in two geographic clusters. *Hum Genet*, 97, 599-603.
- SAWYER, S. L., TIAN, L., KAHKONEN, M., SCHWARTZENTRUBER, J., KIRCHER, M., MAJEWSKI, J., DYMENT, D. A., INNES, A. M., BOYCOTT, K. M., MOREAU, L. A., MOILANEN, J. S. & GREENBERG, R. A. 2014. Biallelic Mutations in BRCA1 Cause a New Fanconi Anemia Subtype. *Cancer Discov*.
- SCHARER, O. D. 2013. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol*, 5, a012609.
- SHECKENBACH, K., BALDUS, S. E., BALZ, V., FREUND, M., PAKROPA, P., SPROLL, C., SCHAFFER, K. L., WAGENMANN, M., SCHIPPER, J. & HANENBERG, H. 2014. RAD51C--a new human cancer susceptibility gene for sporadic squamous cell carcinoma of the head and neck (HNSCC). *Oral Oncol*, 50, 196-9.
- SCHOUTEN, J. P., MCELGUNN, C. J., WAAIJER, R., ZWIJNENBURG, D., DIEPVENS, F. & PALS, G. 2002. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res*, 30, e57.
- SCHUSTER, B., KNIES, K., STOEPKER, C., VELLEUER, E., FRIEDL, R., GOTTWALD-MUHLHAUSER, B., DE WINTER, J. P. & SCHINDLER, D. 2013. Whole exome sequencing reveals uncommon mutations in the recently identified Fanconi anemia gene SLX4/FANCP. *Hum Mutat*, 34, 93-6.
- SCHWARTZ, G. F., HUGHES, K. S., LYNCH, H. T., FABIAN, C. J., FENTIMAN, I. S., ROBSON, M. E., DOMCHEK, S. M., HARTMANN, L. C., HOLLAND, R. & WINCHESTER, D. J. 2008. Proceedings of the international consensus conference on breast cancer risk, genetics, & risk management, April, 2007. *Cancer*, 113, 2627-37.
- SHAH, S., KIM, Y., OSTROVNAYA, I., MURALI, R., SCHRADER, K. A., LACH, F. P., SARREL, K., RAUMURTHY, R., HANSEN, N., ZHANG, L., KIRCHHOFF, T., STADLER, Z., ROBSON, M., VIJAI, J., OFFIT, K. & SMOGORZEWSKA, A. 2013. Assessment of Mutations in Hereditary Breast Cancers. *PLoS One*, 8, e66961.
- SHENDURE, J. & JI, H. 2008. Next-generation DNA sequencing. *Nat Biotechnol*, 26, 1135-45.
- SHENDURE, J., PORRECA, G. J., REPPAS, N. B., LIN, X., MCCUTCHEON, J. P., ROSENBAUM, A. M., WANG, M. D., ZHANG, K., MITRA, R. D. & CHURCH, G. M. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, 309, 1728-32.
- SHIBATA, A., CONRAD, S., BIRRAUX, J., GEUTING, V., BARTON, O., ISMAIL, A., KAKAROUGKAS, A., MEEK, K., TAUCHER-SCHOLZ, G., LOBRICH, M. & JEGGO, P. A. 2011. Factors determining DNA double-strand break repair pathway choice in G2 phase. *EMBO J*, 30, 1079-92.
- SHINOHARA, A. & OGAWA, T. 1998. Stimulation by Rad52 of yeast Rad51-mediated recombination. *Nature*, 391, 404-7.
- SIJBERS, A. M., DE LAAT, W. L., ARIZA, R. R., BIGGERSTAFF, M., WEI, Y. F., MOGGS, J. G., CARTER, K. C., SHELL, B. K., EVANS, E., DE JONG, M. C., RADEMAKERS, S., DE ROOIJ, J., JASPERS, N. G., HOEIJMAKERS, J. H. & WOOD, R. D. 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell*, 86, 811-22.
- SIJBERS, A. M., VAN VOORST VADER, P. C., SNOEK, J. W., RAAMS, A., JASPERS, N. G. & KLEIJER, W. J. 1998. Homozygous R788W point mutation in the XPF gene of a patient with xeroderma pigmentosum and late-onset neurologic disease. *J Invest Dermatol*, 110, 832-6.
- SIMS, A. E., SPITERI, E., SIMS, R. J., 3RD, ARITA, A. G., LACH, F. P., LANDERS, T., WURM, M., FREUND, M., NEVELING, K., HANENBERG, H., AUERBACH, A. D. & HUANG, T. T. 2007. FANCI is a second monoubiquitinated member of the Fanconi anemia pathway. *Nat Struct Mol Biol*, 14, 564-7.

- SINGH, T. R., BAKKER, S. T., AGARWAL, S., JANSEN, M., GRASSMAN, E., GODTHELP, B. C., ALI, A. M., DU, C. H., ROOIMANS, M. A., FAN, Q., WAHENGAM, K., STELTENPOOL, J., ANDREASSEN, P. R., WILLIAMS, D. A., JOENJE, H., DE WINTER, J. P. & MEETEI, A. R. 2009. Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. *Blood*, 114, 174-80.
- SMITH, J., THO, L. M., XU, N. & GILLESPIE, D. A. 2010. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Adv Cancer Res*, 108, 73-112.
- SMOGORZEWSKA, A., DESETTY, R., SAITO, T. T., SCHLABACH, M., LACH, F. P., SOWA, M. E., CLARK, A. B., KUNKEL, T. A., HARPER, J. W., COLAIACOVO, M. P. & ELLEDGE, S. J. 2010. A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Mol Cell*, 39, 36-47.
- SMOGORZEWSKA, A., MATSUOKA, S., VINCIGUERRA, P., MCDONALD, E. R., 3RD, HUROV, K. E., LUO, J., BALLIF, B. A., GYGI, S. P., HOFMANN, K., D'ANDREA, A. D. & ELLEDGE, S. J. 2007. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell*, 129, 289-301.
- SOULIER, J., LEBLANC, T., LARGHERO, J., DASTOT, H., SHIMAMURA, A., GUARDIOLA, P., ESPEROU, H., FERRY, C., JUBERT, C., FEUGEAS, J. P., HENRI, A., TOUBERT, A., SOCIE, G., BARUCHEL, A., SIGAUX, F., D'ANDREA, A. D. & GLUCKMAN, E. 2005. Detection of somatic mosaicism and classification of Fanconi anemia patients by analysis of the FA/BRCA pathway. *Blood*, 105, 1329-36.
- STARESINCIC, L., FAGBEMI, A. F., ENZLIN, J. H., GOURDIN, A. M., WIJGERS, N., DUNAND-SAUTHIER, I., GIGLIA-MARI, G., CLARKSON, S. G., VERMEULEN, W. & SCHARER, O. D. 2009. Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J*, 28, 1111-20.
- STEFANINI, M., KEIJZER, W., DALPRA, L., ELLI, R., PORRO, M. N., NICOLETTI, B. & NUZZO, F. 1980. Differences in the levels of UV repair and in clinical symptoms in two sibs affected by xeroderma pigmentosum. *Hum Genet*, 54, 177-82.
- STOEPKER, C., HAIN, K., SCHUSTER, B., HILHORST-HOFSTEE, Y., ROOIMANS, M. A., STELTENPOOL, J., OOSTRA, A. B., EIRICH, K., KORTHOFF, E. T., NIEUWINT, A. W., JASPERS, N. G., BETTECKEN, T., JOENJE, H., SCHINDLER, D., ROUSE, J. & DE WINTER, J. P. 2011. SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nat Genet*, 43, 138-41.
- STRACKER, T. H. & PETRINI, J. H. 2011. The MRE11 complex: starting from the ends. *Nat Rev Mol Cell Biol*, 12, 90-103.
- STRATHDEE, C. A., DUNCAN, A. M. & BUCHWALD, M. 1992a. Evidence for at least four Fanconi anaemia genes including FACC on chromosome 9. *Nat Genet*, 1, 196-8.
- STRATHDEE, C. A., GAVISH, H., SHANNON, W. R. & BUCHWALD, M. 1992b. Cloning of cDNAs for Fanconi's anaemia by functional complementation. *Nature*, 356, 763-7.
- SU, Y., ORELLI, B., MADIREDDY, A., NIEDERNHOFER, L. J. & SCHARER, O. D. 2012. Multiple DNA binding domains mediate the function of the ERCC1-XPF protein in nucleotide excision repair. *J Biol Chem*, 287, 21846-55.
- SUNG, P. 1994. Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science*, 265, 1241-3.
- SURRALLES, J., JACKSON, S. P., JASIN, M., KASTAN, M. B., WEST, S. C. & JOENJE, H. 2004. Molecular cross-talk among chromosome fragility syndromes. *Genes Dev*, 18, 1359-70.
- SUWAKI, N., KLARE, K. & TARSOUNAS, M. 2011. RAD51 paralogs: roles in DNA damage signalling, recombinational repair and tumorigenesis. *Semin Cell Dev Biol*, 22, 898-905.

- SVENDSEN, J. M., SMOGORZEWSKA, A., SOWA, M. E., O'CONNELL, B. C., GYGI, S. P., ELLEDGE, S. J. & HARPER, J. W. 2009. Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell*, 138, 63-77.
- SY, S. M., HUEN, M. S. & CHEN, J. 2009. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proc Natl Acad Sci U S A*, 106, 7155-60.
- SYMINGTON, L. S. 2014. End resection at double-strand breaks: mechanism and regulation. *Cold Spring Harb Perspect Biol*, 6.
- TAN, D. S., MARCHIO, C. & REIS-FILHO, J. S. 2008. Hereditary breast cancer: from molecular pathology to tailored therapies. *J Clin Pathol*, 61, 1073-82.
- TAN, R., WANG, Y., KLEINSTEIN, S. E., LIU, Y., ZHU, X., GUO, H., JIANG, Q., ALLEN, A. S. & ZHU, M. 2014. An evaluation of copy number variation detection tools from whole-exome sequencing data. *Hum Mutat*, 35, 899-907.
- TANIGUCHI, T., GARCIA-HIGUERA, I., ANDREASSEN, P. R., GREGORY, R. C., GROMPE, M. & D'ANDREA, A. D. 2002a. S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. *Blood*, 100, 2414-20.
- TANIGUCHI, T., GARCIA-HIGUERA, I., XU, B., ANDREASSEN, P. R., GREGORY, R. C., KIM, S. T., LANE, W. S., KASTAN, M. B. & D'ANDREA, A. D. 2002b. Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. *Cell*, 109, 459-72.
- TAYLOR, M. R., SPIREK, M., CHAURASIYA, K. R., WARD, J. D., CARZANIGA, R., YU, X., EGELMAN, E. H., COLLINSON, L. M., RUEDA, D., KREJCI, L. & BOULTON, S. J. 2015. Rad51 Paralogs Remodel Pre-synaptic Rad51 Filaments to Stimulate Homologous Recombination. *Cell*, 162, 271-86.
- THACKER, J. 2005. The RAD51 gene family, genetic instability and cancer. *Cancer Lett*, 219, 125-35.
- THOMPSON, L. H., BUSCH, D. B., BROOKMAN, K., MOONEY, C. L. & GLASER, D. A. 1981. Genetic diversity of UV-sensitive DNA repair mutants of Chinese hamster ovary cells. *Proc Natl Acad Sci U S A*, 78, 3734-7.
- TIAN, M., SHINKURA, R., SHINKURA, N. & ALT, F. W. 2004. Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. *Mol Cell Biol*, 24, 1200-5.
- TIMMERS, C., TANIGUCHI, T., HEJNA, J., REIFSTECK, C., LUCAS, L., BRUUN, D., THAYER, M., COX, B., OLSON, S., D'ANDREA, A. D., MOSES, R. & GROMPE, M. 2001. Positional cloning of a novel Fanconi anemia gene, FANCD2. *Mol Cell*, 7, 241-8.
- TIPPING, A. J., PEARSON, T., MORGAN, N. V., GIBSON, R. A., KUYT, L. P., HAVENGA, C., GLUCKMAN, E., JOENJE, H., DE RAVEL, T., JANSEN, S. & MATHEW, C. G. 2001. Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. *Proc Natl Acad Sci U S A*, 98, 5734-9.
- TOLAR, J., BECKER, P. S., CLAPP, D. W., HANENBERG, H., DE HEREDIA, C. D., KIEM, H. P., NAVARRO, S., QASBA, P., RIO, P., SCHMIDT, M., SEVILLA, J., VERHOEYEN, E., THRASHER, A. J. & BUEREN, J. 2012. Gene therapy for fanconi anemia: one step closer to the clinic. *Hum Gene Ther*, 23, 141-4.
- TRIPSANES, K., FOLKERS, G., AB, E., DAS, D., ODIJK, H., JASPERS, N. G., HOEIJMAKERS, J. H., KAPTEIN, R. & BOELEN, R. 2005. The structure of the human ERCC1/XPF interaction domains reveals a complementary role for the two proteins in nucleotide excision repair. *Structure*, 13, 1849-58.
- TRUJILLO, J. P., MINA, L. B., PUJOL, R., BOGLIOLO, M., ANDRIEUX, J., HOLDER, M., SCHUSTER, B., SCHINDLER, D. & SURRELLES, J. 2012. On the role of FAN1 in Fanconi anemia. *Blood*, 120, 86-9.
- TULPULE, A., LENSCH, M. W., MILLER, J. D., AUSTIN, K., D'ANDREA, A., SCHLAEGER, T. M., SHIMAMURA, A. & DALEY, G. Q. 2010. Knockdown of Fanconi anemia genes in human

- embryonic stem cells reveals early developmental defects in the hematopoietic lineage. *Blood*, 115, 3453-62.
- ULRICH, H. D. 2009a. Regulating post-translational modifications of the eukaryotic replication clamp PCNA. *DNA Repair (Amst)*, 8, 461-469.
- ULRICH, H. D. 2009b. Ubiquitin, SUMO and the maintenance of genome stability. *DNA Repair (Amst)*, 8, 429.
- UTTUKAR, S. M., KLINGEMAN, D. M., LAND, M. L., SCHADT, C. W., DOKTYCZ, M. J., PELLETIER, D. A. & BROWN, S. D. 2014. Evaluation and validation of de novo and hybrid assembly techniques to derive high-quality genome sequences. *Bioinformatics*.
- VAN DIJK, E. L., AUGER, H., JASZCZYSZYN, Y. & THERMES, C. 2014. Ten years of next-generation sequencing technology. *Trends Genet*.
- VAN ZEEBURG, H. J., SNIJDERS, P. J., WU, T., GLUCKMAN, E., SOULIER, J., SURRALLES, J., CASTELLA, M., VAN DER WAL, J. E., WENNERBERG, J., CALIFANO, J., VELLEUER, E., DIETRICH, R., EBELL, W., BLOEMENA, E., JOENJE, H., LEEMANS, C. R. & BRAKENHOFF, R. H. 2008. Clinical and molecular characteristics of squamous cell carcinomas from Fanconi anemia patients. *J Natl Cancer Inst*, 100, 1649-53.
- VAZ, F., HANENBERG, H., SCHUSTER, B., BARKER, K., WIEK, C., ERVEN, V., NEVELING, K., ENDT, D., KESTERTON, I., AUTORE, F., FRATERNALI, F., FREUND, M., HARTMANN, L., GRIMWADE, D., ROBERTS, R. G., SCHAAL, H., MOHAMMED, S., RAHMAN, N., SCHINDLER, D. & MATHEW, C. G. 2010. Mutation of the RAD51C gene in a Fanconi anemia-like disorder. *Nat Genet*, 42, 406-9.
- VENKITARAMAN, A. R. 2002. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*, 108, 171-82.
- VERMEULEN, W., BERGMANN, E., AURIOL, J., RADEMAKERS, S., FRIT, P., APPELDOORN, E., HOEIJMAKERS, J. H. & EGLY, J. M. 2000. Sublimiting concentration of TFIIF transcription/DNA repair factor causes TTD-A trichothiodystrophy disorder. *Nat Genet*, 26, 307-13.
- VIRTS, E. L., JANKOWSKA, A., MACKAY, C., GLAAS, M. F., WIEK, C., KELICH, S. L., LOTTMANN, N., KENNEDY, F. M., MARCHAL, C., LEHNERT, E., SCHARF, R. E., DUFOUR, C., LANCIOTTI, M., FARRUGGIA, P., SANTORO, A., SAVASAN, S., SCHECKENBACH, K., SCHIPPER, J., WAGENMANN, M., LEWIS, T., LEFFAK, M., FARLOW, J. L., FOROUD, T. M., HONISCH, E., NIEDERACHER, D., CHAKRABORTY, S. C., VANCE, G. H., PRUSS, D., TIMMS, K. M., LANCHBURY, J. S., ALPI, A. F. & HANENBERG, H. 2015. AluY-mediated germline deletion, duplication and somatic stem cell reversion in UBE2T defines a new subtype of Fanconi anemia. *Hum Mol Genet*.
- VOLKER, M., MONE, M. J., KARMAKAR, P., VAN HOFFEN, A., SCHUL, W., VERMEULEN, W., HOEIJMAKERS, J. H., VAN DRIEL, R., VAN ZEELAND, A. A. & MULLENDERS, L. H. 2001. Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell*, 8, 213-24.
- WAGNER, J. E., TOLAR, J., LEVRAN, O., SCHOLL, T., DEFFENBAUGH, A., SATAGOPAN, J., BEN-PORAT, L., MAH, K., BATISH, S. D., KUTLER, D. I., MACMILLAN, M. L., HANENBERG, H. & AUERBACH, A. D. 2004. Germline mutations in BRCA2: shared genetic susceptibility to breast cancer, early onset leukemia, and Fanconi anemia. *Blood*, 103, 3226-9.
- WAISFISZ, Q., MORGAN, N. V., SAVINO, M., DE WINTER, J. P., VAN BERKEL, C. G., HOATLIN, M. E., IANZANO, L., GIBSON, R. A., ARWERT, F., SAVOIA, A., MATHEW, C. G., PRONK, J. C. & JOENJE, H. 1999. Spontaneous functional correction of homozygous fanconi anaemia alleles reveals novel mechanistic basis for reverse mosaicism. *Nat Genet*, 22, 379-83.
- WANG, A. T., KIM, T., WAGNER, J. E., CONTI, B. A., LACH, F. P., HUANG, A. L., MOLINA, H., SANBORN, E. M., ZIERHUT, H., CORNES, B. K., ABHYANKAR, A., SOUGNEZ, C., GABRIEL, S. B., AUERBACH,

- A. D., KOWALCZYKOWSKI, S. C. & SMOGORZEWSKA, A. 2015. A Dominant Mutation in Human RAD51 Reveals Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination. *Mol Cell*, 59, 478-90.
- WANG, A. T. & SMOGORZEWSKA, A. 2015. SnapShot: Fanconi anemia and associated proteins. *Cell*, 160, 354-354 e1.
- WANG, B., MATSUOKA, S., BALLIF, B. A., ZHANG, D., SMOGORZEWSKA, A., GYGI, S. P. & ELLEDGE, S. J. 2007. Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science*, 316, 1194-8.
- WANG, R., PERSKY, N. S., YOO, B., OUFELLI, O., SMOGORZEWSKA, A., ELLEDGE, S. J. & PAVLETICH, N. P. 2014. DNA repair. Mechanism of DNA interstrand cross-link processing by repair nuclease FAN1. *Science*, 346, 1127-30.
- WEEDA, G., DONKER, I., DE WIT, J., MORREAU, H., JANSSENS, R., VISSERS, C. J., NIGG, A., VAN STEEG, H., BOOTSMA, D. & HOEIJMAKERS, J. H. 1997. Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr Biol*, 7, 427-39.
- WHITNEY, M., THAYER, M., REIFSTECK, C., OLSON, S., SMITH, L., JAKOBS, P. M., LEACH, R., NAYLOR, S., JOENJE, H. & GROMPE, M. 1995. Microcell mediated chromosome transfer maps the Fanconi anaemia group D gene to chromosome 3p. *Nat Genet*, 11, 341-3.
- WHITNEY, M. A., SAITO, H., JAKOBS, P. M., GIBSON, R. A., MOSES, R. E. & GROMPE, M. 1993. A common mutation in the FACC gene causes Fanconi anaemia in Ashkenazi Jews. *Nat Genet*, 4, 202-5.
- WOOD, R. D. 2010. Mammalian nucleotide excision repair proteins and interstrand crosslink repair. *Environ Mol Mutagen*, 51, 520-6.
- WOOD, R. D., MITCHELL, M., SGOUROS, J. & LINDAHL, T. 2001. Human DNA repair genes. *Science*, 291, 1284-9.
- WOOD, R. D., ROBINS, P. & LINDAHL, T. 1988. Complementation of the xeroderma pigmentosum DNA repair defect in cell-free extracts. *Cell*, 53, 97-106.
- WU, J., ZHANG, X., ZHANG, L., WU, C. Y., REZAEIAN, A. H., CHAN, C. H., LI, J. M., WANG, J., GAO, Y., HAN, F., JEONG, Y. S., YUAN, X., KHANNA, K. K., JIN, J., ZENG, Y. X. & LIN, H. K. 2012. Skp2 E3 ligase integrates ATM activation and homologous recombination repair by ubiquitinating NBS1. *Mol Cell*, 46, 351-61.
- XIA, B., DORSMAN, J. C., AMEZIANE, N., DE VRIES, Y., ROOIMANS, M. A., SHENG, Q., PALS, G., ERRAMI, A., GLUCKMAN, E., LLERA, J., WANG, W., LIVINGSTON, D. M., JOENJE, H. & DE WINTER, J. P. 2007. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat Genet*, 39, 159-61.
- XIA, B., SHENG, Q., NAKANISHI, K., OHASHI, A., WU, J., CHRIST, N., LIU, X., JASIN, M., COUCH, F. J. & LIVINGSTON, D. M. 2006. Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell*, 22, 719-29.
- YAGASAKI, H., HAMANOUE, S., ODA, T., NAKAHATA, T., ASANO, S. & YAMASHITA, T. 2004. Identification and characterization of novel mutations of the major Fanconi anemia gene FANCA in the Japanese population. *Hum Mutat*, 24, 481-90.
- YOKOYAMA, A., YOKOYAMA, T., BROOKS, P. J., MIZUKAMI, T., MATSUI, T., KIMURA, M., MATSUSHITA, S., HIGUCHI, S. & MARUYAMA, K. 2014. Macrocytosis, macrocytic anemia, and genetic polymorphisms of alcohol dehydrogenase-1B and aldehyde dehydrogenase-2 in Japanese alcoholic men. *Alcohol Clin Exp Res*, 38, 1237-46.
- YONETANI, Y., HOCHEGGER, H., SONODA, E., SHINYA, S., YOSHIKAWA, H., TAKEDA, S. & YAMAZOE, M. 2005. Differential and collaborative actions of Rad51 paralog proteins in cellular response to DNA damage. *Nucleic Acids Res*, 33, 4544-52.

- YOON, J. H., LEE, C. S., O'CONNOR, T. R., YASUI, A. & PFEIFER, G. P. 2000. The DNA damage spectrum produced by simulated sunlight. *J Mol Biol*, 299, 681-93.
- YOUSSOUFIAN, H. 1996. Natural gene therapy and the Darwinian legacy. *Nat Genet*, 13, 255-6.
- YUN, M. H. & HIOM, K. 2009. CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature*, 459, 460-3.
- ZHANG, F., MA, J., WU, J., YE, L., CAI, H., XIA, B. & YU, X. 2009. PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Curr Biol*, 19, 524-9.
- ZHANG, J., BARBARO, P., GUO, Y., ALODAIB, A., LI, J., GOLD, W., ADES, L., KEATING, B. J., XU, X., TEO, J., HAKONARSON, H. & CHRISTODOULOU, J. 2015. Utility of next-generation sequencing technologies for the efficient genetic resolution of haematological disorders. *Clin Genet*.
- ZHAO, Q., XUE, X., LONGERICH, S., SUNG, P. & XIONG, Y. 2014. Structural insights into 5' flap DNA unwinding and incision by the human FAN1 dimer. *Nat Commun*, 5, 5726.
- ZHAO, W., VAITHIYALINGAM, S., SAN FILIPPO, J., MARANON, D. G., JIMENEZ-SAINZ, J., FONTENAY, G. V., KWON, Y., LEUNG, S. G., LU, L., JENSEN, R. B., CHAZIN, W. J., WIESE, C. & SUNG, P. 2015. Promotion of BRCA2-Dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry. *Mol Cell*, 59, 176-87.
- ZHOU, W., OTTO, E. A., CLUCKEY, A., AIRIK, R., HURD, T. W., CHAKI, M., DIAZ, K., LACH, F. P., BENNETT, G. R., GEE, H. Y., GHOSH, A. K., NATARAJAN, S., THONGTHIP, S., VETURI, U., ALLEN, S. J., JANSSEN, S., RAMASWAMI, G., DIXON, J., BURKHALTER, F., SPOENDLIN, M., MOCH, H., MIHATSCH, M. J., VERINE, J., READE, R., SOLIMAN, H., GODIN, M., KISS, D., MONGA, G., MAZZUCCO, G., AMANN, K., ARTUNC, F., NEWLAND, R. C., WIECH, T., ZSCHIEDRICH, S., HUBER, T. B., FRIEDL, A., SLAATS, G. G., JOLES, J. A., GOLDSCHMEDING, R., WASHBURN, J., GILES, R. H., LEVY, S., SMOGORZEWSKA, A. & HILDEBRANDT, F. 2012. FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. *Nat Genet*, 44, 910-5.

ANEXOS

IX. ANEXOS

Anexo 1:

Trujillo JP, Mina LB, Pujol R, **Bogliolo M**, Andrieux J, Holder M, Schuster B, Schindler D, Surrallés J. **On the role of FAN1 in Fanconi anemia.** Blood. 2012 Jul 5;120(1):86-9.

Brief report

On the role of FAN1 in Fanconi anemia

*Juan P. Trujillo,^{1,2} *Leonardo B. Mina,^{1,2} Roser Pujol,^{1,2} Massimo Bogliolo,^{1,2} Joris Andrieux,³ Muriel Holder,⁴ Beatrice Schuster,⁵ Detlev Schindler,⁵ and Jordi Surrallés^{1,2}

¹Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain;

²Centre for Biomedical Network Research on Rare Diseases, Instituto de Salud Carlos III, Bellaterra, Barcelona, Spain; ³Institut de Génétique Médicale, Hôpital Jeanne de Flandre, Centre Hospitalier Régional Universitaire de Lille, Lille, France; ⁴Génétique Clinique, Hôpital Jeanne de Flandre, Centre Hospitalier Régional Universitaire de Lille, Lille, France; and ⁵Department of Human Genetics, University of Würzburg, Würzburg, Germany

Fanconi anemia (FA) is a rare bone marrow failure disorder with defective DNA interstrand crosslink repair. Still, there are FA patients without mutations in any of the 15 genes individually underlying the disease. A candidate protein for those patients, FA nuclease 1 (FAN1), whose gene is located at chromosome 15q13.3, is recruited to stalled replication forks by binding to monoubiquitinated FANCD2 and is required for interstrand crosslink

repair, suggesting that mutation of FAN1 may cause FA. Here we studied clinical, cellular, and genetic features in 4 patients carrying a homozygous 15q13.3 microdeletion, including FAN1 and 6 additional genes. Biallelic deletion of the entire FAN1 gene was confirmed by failure of 3'- and 5'-PCR amplification. Western blot analysis failed to show FAN1 protein in the patients' cell lines. Chromosome fragility was normal in all 4 FAN1-deficient pa-

tients, although their cells showed mild sensitivity to mitomycin C in terms of cell survival and G₂ phase arrest, dissimilar in degree to FA cells. Clinically, there were no symptoms pointing the way to FA. Our results suggest that FAN1 has a minor role in interstrand crosslink repair compared with true FA genes and exclude FAN1 as a novel FA gene. (*Blood*. 2012;120(1):86-89)

Introduction

Fanconi anemia (FA) is characterized by chromosome breakage, congenital malformations, pancytopenia, and cancer susceptibility.¹ FA is a rare disease with a carrier frequency of 1:65 to 1:209.^{2,3} FA cells are hypersensitive to DNA interstrand crosslinking (ICL) drugs, such as mitomycin C (MMC) and diepoxybutane (DEB), and the diagnosis relies on an excess chromosome fragility after in vitro exposing patients' cells to these agents. There are at least 15 independent FA subtypes, each resulting from mutation of a distinct FA gene.⁴⁻⁷ However, a minority of FA patients remain unassigned, suggesting the existence of additional FA genes. Recently, 4 groups reported that FA nuclease 1 (FAN1) is a good candidate for a novel FA gene.⁸⁻¹¹ The reason is that FAN1 is recruited to stalled replication forks by binding to monoubiquitinated FANCD2, and its nuclease activity is required for ICL repair. Transient depletion of FAN1 in human transformed fibroblasts led to increased MMC-induced chromosome breakage rates. Consequently, all 4 groups suggested that FAN1 mutations may cause FA.⁸⁻¹¹

FAN1 maps to 15q13.3. Heterozygous 15q13.3 microdeletion has been associated with a variety of symptoms, including mental retardation, epilepsy, psychiatric disease, autism spectrum disorders, muscular hypotonia, and dysmorphic facial features. Penetration of the microdeletion disorder is variable and encompasses severely affected patients to normal persons.¹² Apart from FAN1, 6 additional genes are located in 15q13.3 (ARHGAP11B, MTMR10, TRPM1, KLF13, OTUD7A, and CHRNA7). Here we studied 4 patients with homozygous 15q13.3 microdeletion^{12,13} to clarify whether lack of FAN1 may lead to FA.

Methods

Clinical features and blood samples were obtained from 4 homozygous 15q13.3 microdeletion patients (MD1-MD4) all previously diagnosed by array comparative genomic hybridization and quantitative PCR. Two of these patients (MD1 and MD2) have been mentioned before.^{12,13} Lack of FAN1 was confirmed at the gene level by PCR and at the protein level by Western blotting. The PCR primers used to amplify the 3' and 5' flanking regions of FAN1 were as follows: ex1 forward, 5'AGGGTTGTCTCCTCGT-TACAGGA3'; ex1 reverse, 5'GCTGAATCACTTTGGCCAGG3'; ex15 forward, 5'CTTCCTAAAACCTGCTGGAGG3'; and ex15 reverse, 5'AATGTACTGACCGTGTGCTCA3'. PCR, Western blot analysis, survival assays, and chromosome breakage assays were performed as described elsewhere.^{3,14-18} FAN1-monospecific antibody was kindly provided by Dr John Rouse (Dundee, United Kingdom) and used at 1:500 dilution. A total of 27 genetically unassigned FA cell lines had previously been excluded from belonging to any of the reported 15 FA complementation groups. This study was ethically approved by the Universitat Autònoma de Barcelona Institutional Review Board. Informed consent was obtained from all families in accordance with the Declaration of Helsinki.

Results and discussion

Study of FA candidate genes may enable the final classification of unassigned FA patients. Four recent studies have proposed FAN1 as a putative FA gene.⁸⁻¹¹ Here we studied 4 patients (MD1-MD4) with homozygous 15q13.3 microdeletion to clarify whether FAN1 deficiency leads to features consistent with FA. Two of these

Submitted March 30, 2012; accepted May 15, 2012. Prepublished online as *Blood* First Edition paper, May 18, 2012; DOI 10.1182/blood-2012-04-420604.

*J.P.T. and L.B.M. contributed equally to this study.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2012 by The American Society of Hematology

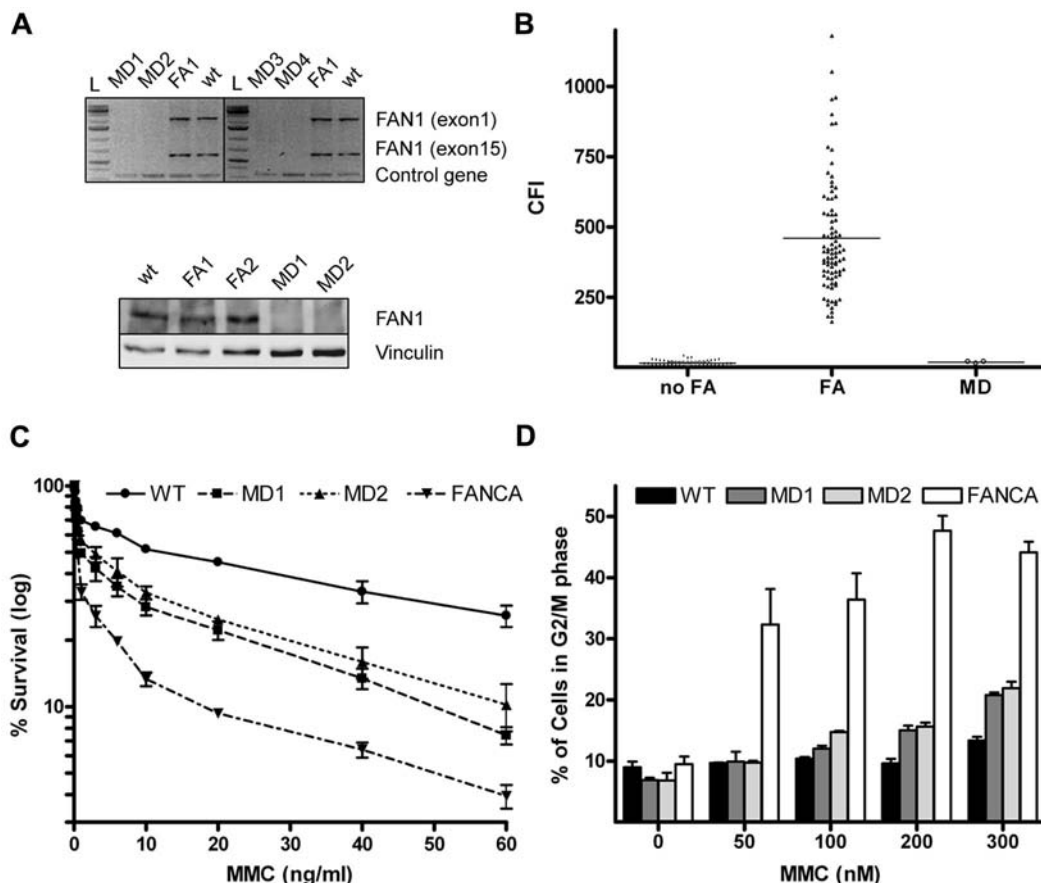


Figure 1. Absence of FAN1 and FA cellular phenotype in MD patients. (A) Fragments amplified by PCR corresponding to exons 1 and 15 of the *FAN1* gene were observed using DNA templates from a healthy person (WT) and a FANCA-deficient FA patient included as controls, whereas they were absent when templates from microdeletion patients (MD1 to MD4) were used, confirming the biallelic deletion of *FAN1* in the MD patients (top panel). Immunoblotting against FAN1 protein revealed a FAN1-specific band in WT and FA (FA1 and FA2) LCL that failed to be detected in MD1 and MD2. (bottom panel). (B) Dot plot of CFI showing individual values and average (solid line) of DEB-induced chromosome breakage from non-FA ($n = 56$), FA ($n = 90$, excluding mosaics) and MD ($n = 4$) persons. The CFI values of all MD patients ranged within the non-FA population. (C) Mild sensitivity of MD LCL to MMC on survival assay. The graph shows intermediate sensitivity to MMC of MD1 and MD2 compared with the highly sensitive FA (FANCA) cell line. (D) Near-normal sensitivity of MD LCL to MMC on cell cycle analysis. The graph plots the percentage of cells in G2/M phase after exposure to increasing concentrations of MMC for 72h. A WT and an FANCA cell lines were included as controls.

patients (MD1 and MD2) have previously been mentioned in unrelated reports^{12,13} and the other 2 are newly recognized siblings, detected by array comparative genomic hybridization (supplemental Figure 1, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article). Considering that 15q13.3 microdeletion may have 3 different extensions,¹⁹ we aimed to corroborate homozygous absence of the *FAN1* gene by failure to PCR amplify its first (exon 1) and last (exon 15) exon from genomic DNA. DNA of 2 additional lymphoblastoid cell lines (LCLs), from a normal person and from a FANCA-deficient patient, served as controls. As shown in Figure 1A top panel, FAN1 PCR products are absent in all MD patients, confirming that all 4 MD patients have biallelic deletion of the entire *FAN1* gene. Moreover, lack of FAN1 expression was confirmed by FAN1 immunoblotting. Clearly, the FAN1-specific band was missing in the 2 available LCLs from MD patients (MD1 and MD2), whereas FAN1 was readily detected in the control LCLs (Figure 1A bottom panel).

To check whether FAN1 deficiency leads to DEB-induced chromosome fragility,^{20,21} we performed DEB tests on an LCL from patient MD1 and on blood T cells from patients MD2 to MD4. Chromosome breakage rates were quantified with the recently described chromosome fragility index (CFI)¹⁷ and the results

compared with our historical database.¹⁷ Clearly, the CFI of all MD patients fell into the range of the non-FA group (Figure 1B). Similar results were obtained with MMC (data not shown).

We next tested the survival of the 2 available MD LCLs in response to MMC. MD1, MD2, a wild-type and a FANCA LCL were challenged with 0 to 100 ng/mL of MMC. Based on LD₅₀ values, the MD cell lines showed mild sensitivity to MMC: whereas FANCA-deficient cells were more than 30-fold more sensitive to MMC than WT cells, MD1 and MD2 cells were, on average, 5-fold more sensitive to MMC than WT cells (Figure 1C). Silencing of the *FAN1* gene by siRNA was previously shown to impair ICL repair, leading to hypersensitivity of cells to ICL. However, this hypersensitivity was also intermediate compared with mRNA depletion of authentic FA genes, such as *FANCA*, *FANCD2*, or *FANCF*.⁸⁻¹⁰ This set of data suggests that the cellular response of FAN1-deficient cells to MMC is not fully functional but not impaired as in FA.

To further study the FA pathway in the absence of FAN1, cell cycle distributions of FAN1-deficient cell lines were analyzed by flow cytometry.²² Exposure to increasing concentrations of MMC for 72 hours resulted in G₂ arrest at low MMC concentrations in FA-A LCL, whereas G₂ arrest was very mild in the MD samples (Figure 1D), compatible with the mild sensitivity to MMC shown

Table 1. Patient characteristics

Patient no.	Age, y	Nationality	Clinical features	Hematology	Chromosome fragility	Sensitivity to MMC	G ₂ /M block	Reference
MD1	11	United States	Visual impairment, hypotonia, areflexia, absent language, epilepsy, microsomy, and microcephaly	Normal	Negative	Mild	Mild	13
MD2	6	France	Hypotonia, severe developmental delay; rod–cone dystrophy, epilepsy, and autistic features	Normal	Negative	Mild	Mild	12
MD3	1	France	Severe developmental delay, visual impairment, microsomy, and microcephaly	Normal	Negative	ND	ND	Present study
MD4	3	France	Severe developmental delay, absent language, visual impairment, microsomy, and microcephaly	Normal	Negative	ND	ND	Present study

MMC indicates mitomycin C; and MD, microdeletion.

before. These results are consistent with a recent report on Δ FAN1-DT40 cells showing that FAN1 protects cells against ICL agents in a pathway, which is not epistatic with the FA pathway and that FAN1 assumes in the processing of ICL only a secondary role or functions independently of the FA pathway.²³ We finally analyzed FAN1 protein expression levels in 27 cell line from unassigned FA patients by Western blotting. All of the unassigned FA cell lines expressed FAN1 protein at control levels, suggesting that none of these patients had major deficiency of this protein (supplemental Figure 2).

To assess the hematologic impact of FAN1 deficiency, we obtained clinical data and hemograms of MD2, MD3, and MD4. Normal hematology had earlier been reported for MD1.¹² As shown in Table 1, MD patients do not present with anemia, bone marrow failure, skin pigmentation anomalies, or FA-typical malformations, such as skeletal abnormalities of the upper limbs. Three of the MD patients (MD1, MD3, and MD4) showed microsomy and microcephaly, which is often seen in FA patients but also in other syndromes with defective processing of stalled replication forks, such as Seckle and Bloom syndromes and can be regarded as common symptoms of patients with DNA repair defects.²⁴ Yet we cannot conclude for certain that microcephaly and microsomy found in MD patients are caused by FAN1 deficiency because 6 additional genes are included in the 15q13.3 region. However, it is tempting to speculate that this is the case as FAN1 directly interacts with FANCD2, and 90% of patients with FANCD2 mutations have microcephaly.¹⁵

Even though LCLs with total FAN1 deficiency reveal mild sensitivity to MMC on some assays, normal expression of FAN1 in 27 unassigned FA cell lines, the lack of DEB- or MMC-induced chromosome fragility, and the absence of hematologic defects or FA-archetypal malformations exclude FAN1 as being an FA gene.

References

- Kutler DI, Singh B, Satagopan J, et al. A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood*. 2003;101(4):1249-1256.
- Rosenberg PS, Tamary H, Alter BP. 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*. 2011;155A(8):1877-1883.
- Callén E, Casado JA, Tischkowitz MD, et al. A common founder mutation in FANCA underlies the world's highest prevalence of Fanconi anemia in Gypsy families from Spain. *Blood*. 2005;105(5):1946-1949.
- Wang W. Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat Rev Genet*. 2007;8(10):735-748.
- Meindl A, Hellebrand H, Wiek C, et al. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat Genet*. 2010;42(5):410-414.
- Meetei AR, Medhurst AL, Ling C, et al. A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet*. 2005;37(9):958-963.
- Kim Y, Lach FP, Desetty R, Hanenberg H, Auerbach AD, Smogorzewska A. Mutations of the SLX4 gene in Fanconi anemia. *Nat Genet*. 2011;43(2):142-146.
- Kratz K, Schopf B, Kaden S, et al. Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. *Cell*. 2010;142(1):77-88.
- Liu T, Ghosal G, Yuan J, Chen J, Huang J. FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. *Science*. 2010;329(5992):693-696.
- MacKay C, Declais AC, Lundin C, et al. Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell*. 2010;142(1):65-76.
- Smogorzewska A, Desetty R, Saito TT, et al. A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA

Acknowledgments

The authors thank Dr B. C. Bittel (University of Missouri–Kansas City School of Medicine, Kansas City, MO) for providing an LCL from patient MD1 and Dr J. Rouse (University of Dundee, Dundee, United Kingdom) for sharing his anti-FAN1 antibody.

The laboratory of J.S. was supported by the Generalitat de Catalunya (SGR0489-2009), the Institut Català de Recerca i Estudis Avançats-Academia award, the Spanish Ministry of Science and Innovation (projects CB06/07/0023, and SAF2009-11936), and the European Regional Development Funds. Centre for Biomedical Network Research on Rare Diseases is an initiative of the Instituto de Salud Carlos III.

Authorship

Contribution: J.P.T., L.B.M., R.P., M.B., and B.S. performed experiments and helped write the manuscript; J.A. and M.H. provided essential research materials and clinical data and performed experiments; D.S. designed experiments and provided essential research materials; and J.S. coordinated and supervised the study, designed experiments, and wrote the paper with the help of L.B.M.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Jordi Surrallés, Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Campus de Bellaterra S/N, 08193, Bellaterra, Barcelona, Spain; e-mail: jordi.surralles@uab.es.

- interstrand crosslink repair. *Mol Cell*. 2010;39(1):36-47.
12. Masurel-Paulet A, Andrieux J, Callier P, et al. Delineation of 15q13.3 microdeletions. *Clin Genet*. 2010;78(2):149-161.
 13. Lepichon JB, Bittel DC, Graf WD, Yu S. A 15q13.3 homozygous microdeletion associated with a severe neurodevelopmental disorder suggests putative functions of the TRPM1, CHRNA7, and other homozygously deleted genes. *Am J Med Genet A*. 2010;152A(5):1300-1304.
 14. Bogliolo M, Lyakhovich A, Callen E, et al. Histone H2AX and Fanconi anemia FANCD2 function in the same pathway to maintain chromosome stability. *EMBO J*. 2007;26(5):1340-1351.
 15. Kalb R, Neveling K, Hoehn H, et al. Hypomorphic mutations in the gene encoding a key Fanconi anemia protein, FANCD2, sustain a significant group of FA-D2 patients with severe phenotype. *Am J Hum Genet*. 2007;80(5):895-910.
 16. Raya A, Rodriguez-Piza I, Guenechea G, et al. Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature*. 2009;460(7251):53-59.
 17. Castella M, Pujol R, Callen E, et al. Chromosome fragility in patients with Fanconi anaemia: diagnostic implications and clinical impact. *J Med Genet*. 2011;48(4):242-250.
 18. Castella M, Pujol R, Callen E, et al. Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. *Blood*. 2011;117(14):3759-3769.
 19. Liao J, DeWard SJ, Madan-Khetarpal S, Surti U, Hu J. A small homozygous microdeletion of 15q13.3 including the CHRNA7 gene in a girl with a spectrum of severe neurodevelopmental features. *Am J Med Genet A*. 2011;155A(11):2795-2800.
 20. Antonio Casado J, Callen E, Jacome A, et al. A comprehensive strategy for the subtyping of patients with Fanconi anaemia: conclusions from the Spanish Fanconi Anemia Research Network. *J Med Genet*. 2007;44(4):241-249.
 21. Auerbach AD. Diagnosis of Fanconi anemia by diepoxybutane analysis. *Curr Protoc Hum Genet*. 2003;Chapter 8:Unit 8.7.
 22. Ceccaldi R, Briot D, Larghero J, et al. Spontaneous abrogation of the G2DNA damage checkpoint has clinical benefits but promotes leukemogenesis in Fanconi anemia patients. *J Clin Invest*. 2011;121(1):184-194.
 23. Yoshikiyo K, Kratz K, Hirota K, et al. KIAA1018/FAN1 nuclease protects cells against genomic instability induced by interstrand cross-linking agents. *Proc Natl Acad Sci U S A*. 2010;107(50):21553-21557.
 24. O'Driscoll M, Jeggo PA. The role of the DNA damage response pathways in brain development and microcephaly: insight from human disorders. *DNA Repair (Amst)*. 2008;7(7):1039-1050.

Anexo 2

Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A, Trujillo JP, Minguillón J, Ramírez MJ, Pujol R, Casado JA, Baños R, Rio P, Knies K, Zúñiga S, Benítez J, Bueren JA, Jaspers NG, Schärer OD, de Winter JP, Schindler D, Surrallés J. **Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia.** Am J Hum Genet. 2013 May 2;92(5):800-6.

Mutations in *ERCC4*, Encoding the DNA-Repair Endonuclease XPF, Cause Fanconi Anemia

Massimo Bogliolo,^{1,2,10} Beatrice Schuster,^{3,10} Chantal Stoeperker,⁴ Burak Derkunt,⁵ Yan Su,⁵ Anja Raams,⁶ Juan P. Trujillo,¹ Jordi Minguillón,¹ María J. Ramírez,^{1,2} Roser Pujol,^{1,2} José A. Casado,^{2,7} Rocío Baños,^{2,7} Paula Rio,^{2,7} Kerstin Knies,³ Sheila Zúñiga,⁸ Javier Benítez,^{2,9} Juan A. Bueren,^{2,7} Nicolaas G.J. Jaspers,⁶ Orlando D. Schärer,⁵ Johan P. de Winter,⁴ Detlev Schindler,^{3,*} and Jordi Surrallés^{1,2,*}

Fanconi anemia (FA) is a rare genomic instability disorder characterized by progressive bone marrow failure and predisposition to cancer. FA-associated gene products are involved in the repair of DNA interstrand crosslinks (ICLs). Fifteen FA-associated genes have been identified, but the genetic basis in some individuals still remains unresolved. Here, we used whole-exome and Sanger sequencing on DNA of unclassified FA individuals and discovered biallelic germline mutations in *ERCC4* (*XPF*), a structure-specific nuclease-encoding gene previously connected to xeroderma pigmentosum and segmental XFE progeroid syndrome. Genetic reversion and wild-type *ERCC4* cDNA complemented the phenotype of the FA cell lines, providing genetic evidence that mutations in *ERCC4* cause this FA subtype. Further biochemical and functional analysis demonstrated that the identified FA-causing *ERCC4* mutations strongly disrupt the function of XPF in DNA ICL repair without severely compromising nucleotide excision repair. Our data show that depending on the type of *ERCC4* mutation and the resulting balance between both DNA repair activities, individuals present with one of the three clinically distinct disorders, highlighting the multifunctional nature of the XPF endonuclease in genome stability and human disease.

Fanconi anemia (FA) is characterized by bone marrow failure (BMF), congenital malformations, hypersensitivity to DNA interstrand crosslink (ICL)-inducing agents, chromosome fragility, and a high susceptibility to cancer. Since the discovery of the first FA-associated gene 20 years ago, all together, 15 genes associated with FA have been identified; these include *FANCA*, *FANCB*, *FANCC*, *FANCD1* (*BRCA2*), *FANCD2*, *FANCE*, *FANCF*, *FANCG* (*XRCC9*), *FANCI*, *FANCL* (*BRIP1*), *FANCL* (*PHF9*), *FANCM*, *FANCN* (*PALB2*), *FANCO* (*RAD51C*), and *FANCP* (*SLX4*)^{1,2} (MIM 227650, 300514, 227645, 605724, 227646, 600901, 603467, 614082, 609053, 609054, 614083, 614087, 610832, 613390, and 613951, respectively). Studies to unravel the genetic basis of this rare disorder uncovered a genome-maintenance pathway that protects dividing cells against replication-blocking DNA lesions. To identify additional FA-associated genes, we used the SOLiD 4 platform for whole-exome sequencing on peripheral-blood DNA from a Spanish FA individual (FA104) who was previously excluded from all known FA complementation groups (this study was approved by the Institutional Committee on Ethical Research in Human Samples, and proper informed consent was obtained). FA104 was born to unrelated parents and was diagnosed neonatally with a malformative syndrome suggestive of FA, the symptoms of which included bilateral absent thumbs, microsomy, esophageal atresia, a ventrally translocated anus, and dysplastic and

low-set ears. She did not show any dermatological abnormality such as skin hyperpigmentation, photosensitivity, sunlight-induced scarring, or atrophy. FA104 developed BMF at the age of 2 years and died as a result of a hemorrhagic shock after bone marrow transplantation at the age of 4 years. A positive chromosome-breakage test unambiguously confirmed the FA diagnosis: 92% of the cells showed on average 4.4 diepoxybutane (DEB)-induced breaks. Lymphoblasts from this individual were hypersensitive to mitomycin-C (MMC) and melphalan but were insensitive to the topoisomerase I inhibitor camptothecin and the PARP inhibitor KU58948 (data not shown) and showed normal *FANCD2* monoubiquitination and *RAD51* focus formation.³ This suggests a defect downstream within the FA pathway, which does not involve homologous recombination. On the basis of a recessive mode of inheritance, exome sequencing identified 17 candidate disease genes for FA104 (Table S1, available online); of these, *ERCC4* (MIM 133520; also known as *XPF*) immediately caught our attention given the involvement of the XPF-ERCC1-structure-specific nuclease in ICL repair.⁴ Both *ERCC4* mutations were predicted to be pathogenic: a 5 bp deletion in exon 8 (c.1484_1488delCTCAA) was predicted to lead to a frameshift and a premature stop codon (p.Thr495Asnfs*6), and a missense mutation in exon 11 (c.2065C>A [p.Arg689Ser]; RefSeq accession numbers NG_011442.1, NM_005236.2, and NP_005227.1) was

¹Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain; ²Centre for Biomedical Network Research on Rare Diseases, Instituto de Salud Carlos III, Bellaterra, 08193 Barcelona, Spain; ³Department of Human Genetics, University of Wurzburg, D-97074 Wurzburg, Germany; ⁴Department of Clinical Genetics, VU University Medical Center, 1081 BT Amsterdam, the Netherlands; ⁵Department of Pharmacological Sciences and Chemistry, Stony Brook University, Stony Brook, NY 11794-3400, USA; ⁶Department of Genetics, Erasmus University Medical Center, 3015 GE Rotterdam, the Netherlands; ⁷Hematopoiesis and Gene Therapy Division, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, 28040 Madrid, Spain; ⁸Department of Bioinformatics, Sistemas Genómicos SL, 46980 Valencia, Spain; ⁹Human Genetics Group, Spanish National Cancer Center, Centro Nacional de Investigaciones Oncológicas, E-28029 Madrid, Spain

¹⁰These authors contributed equally to this work

*Correspondence: jordi.surralles@uab.es (J.S.), schindler@biozentrum.uni-wuerzburg.de (D.S.)

<http://dx.doi.org/10.1016/j.ajhg.2013.04.002>. ©2013 by The American Society of Human Genetics. All rights reserved.

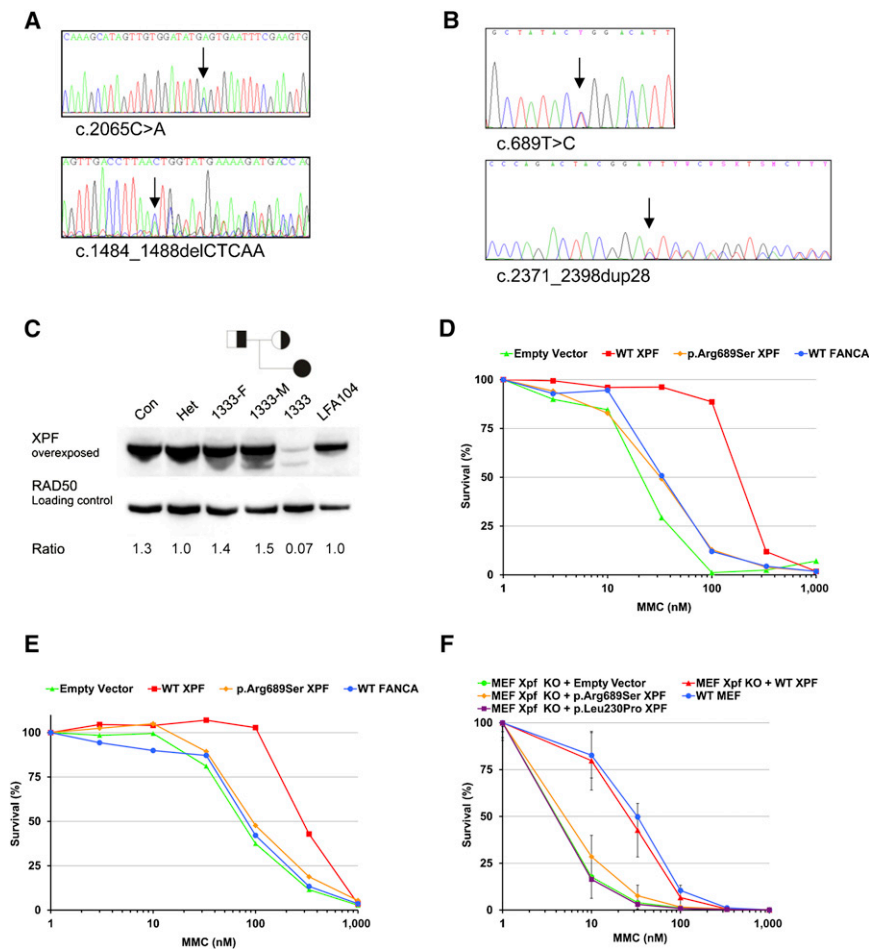


Figure 1. ERCC4 Mutations and XPF Deficiency in FA Individuals

(A) Sequence analysis of blood DNA from FA104 revealed a missense mutation in exon 11 (c.2065C>A [p. Arg689Ser]) (upper panel) and a 5 bp deletion in exon 8 leading to a frameshift and premature termination of translation (c.1484_1488delCTCAA [p. Thr495Asnfs*6]) (lower panel).

(B) Sequence analysis of blood DNA from 1333 revealed a missense mutation in exon 4 (c.689T>C [p.Leu230Pro]) (upper panel) and a 28 bp duplication in exon 11 (lower panel) leading to a frameshift and a premature stop codon (c.2371_2398dup28 [p.Ile800Thrfs*24]).

(C) Immunoblot analysis showing XPF expression in lymphoblasts from 1333 and FA104. Lymphoblasts from a healthy individual (Con), the parents of 1333 (1333-F and 1333-M), and an unrelated ERCC4 mutation carrier (Het) were used as controls. XPF levels are expressed as a ratio relative to the loading control (RAD50).

(D) Genetic complementation of MMC sensitivity in FA104 lymphoblasts by wild-type XPF, but not by p.Arg689Ser altered XPF. Site-directed mutagenesis was used for introducing point mutations into the pWPKL-XPF hemagglutinin (HA)-tagged plasmid with the QuickChange method (Stratagene) as described.⁵ Lentiviral supernatant production and transduction were done as previously described,⁶ and cells were grown for 10 days in the presence of MMC. Data represent a typical result of at least three independent experiments.

(E) Genetic complementation of MMC sensitivity of 1333 lymphoblasts by wild-type XPF (experiments were performed as in D).

(F) MMC-induced growth inhibition of *Ercc4*-knockout MEFs transduced with lentiviral particles coding for GFP (negative control vector), wild-type XPF, and p.Arg689Ser and p.Leu230Pro altered XPF. Data represent means and SD of at least three independent experiments.

predicted to change a highly conserved arginine within the nuclease active site of XPF. Sanger sequencing on blood DNA confirmed these mutations (Figure 1A) and their correct segregation (data not shown). In MMC-resistant FA104 lymphoblasts (FA104R) obtained after long-term exposure to a low dose of MMC, we detected a mutation that restored the ERCC4 reading frame (Figure S1A), supporting the notion that MMC sensitivity is due to ERCC4 mutations. Consistently, XPF levels were reduced in FA104 lymphoblasts but were normalized in the reverted FA104R lymphoblasts (Figure S1B). Immunoblotting did not detect a truncated XPF, indicating that only the p.Arg689Ser altered XPF was present in the FA104 cell line.

Sanger sequencing on 18 unclassified FA individuals from Germany revealed biallelic ERCC4 mutations in another individual (1333). Individual 1333 was born in 2002 and was unambiguously diagnosed with FA at the age of 5 years as a result of multiple FA-related features, such as perinatal growth retardation, short stature, pronounced microcephaly, café-au-lait spots, an ostium-primum defect, biliary atresia with fibrosis of the liver, BMF,

and a positive chromosome-fragility test (0.2, 6.7, and 9.4 breaks per cell at 0, 50, and 100 ng/ml MMC, respectively). Individual 1333 is redheaded and has pale skin color, but no spontaneous or UV-light-induced skin lesions were reported at the age of 10 years. Similar to those of FA104, lymphoblasts from individual 1333 were normal with regard to FANCD2 monoubiquitination and RAD51 focus formation and were sensitive to MMC and melphalan but insensitive to the topoisomerase I inhibitor camptothecin and to the PARP inhibitor KU58948 (data not shown). Individual 1333 carries a 28 bp duplication in exon 11 of the maternal allele (c.2371_2398dup28 [p.Ile800Thrfs*24]; Figure 1B), and this duplication is predicted to result in a truncated XPF that lacks the double helix-hairpin-helix (HhH₂) domain involved in heterodimerization with ERCC1 and DNA binding.⁷ The paternal allele contains a missense mutation that changes a highly conserved amino acid residue within the helicase-like domain (c.689T>C [p. Leu230Pro]; Figure 1B). Immunoblot analysis showed that a missense altered XPF and a truncated 90–95 kDa XPF are present at very low levels

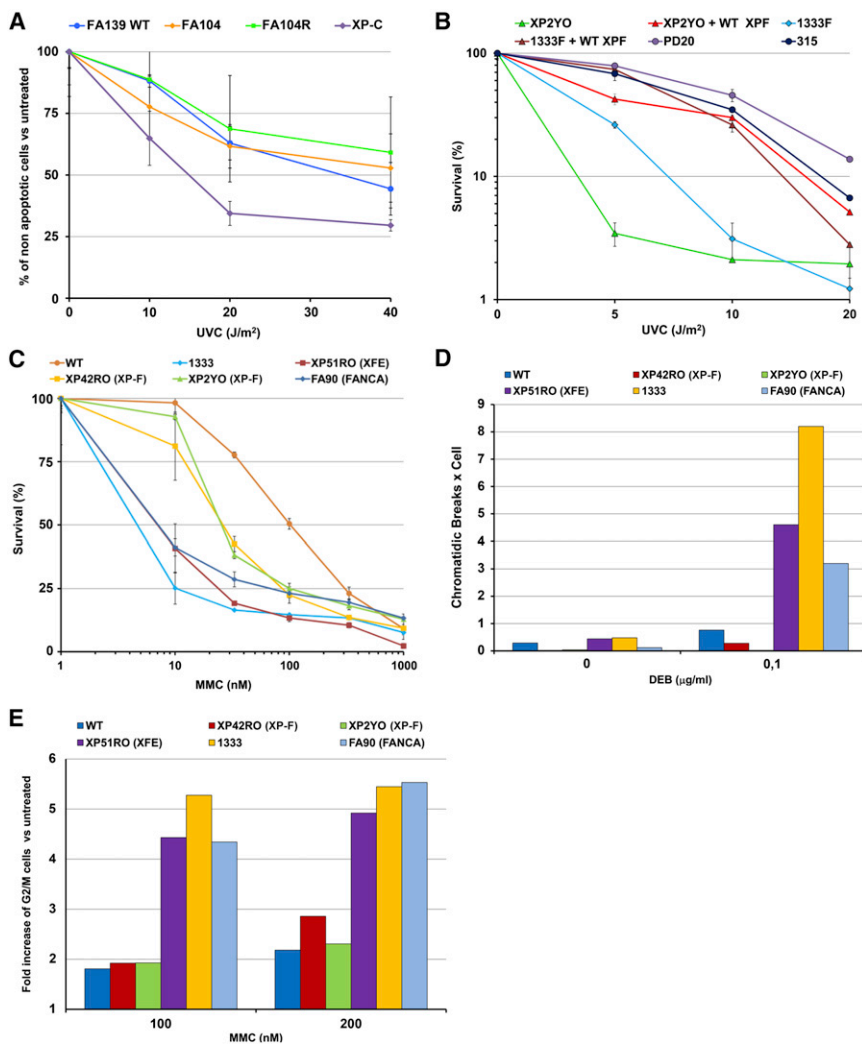


Figure 2. UV-Light and ICL Sensitivities of *ERCC4* Mutants Leading to FA

(A) UVC-light-induced apoptosis in FA104 lymphoblasts. Cells were analyzed for UVC-light-induced apoptosis 24 hr after irradiation with the use of the Annexin-V-FLUOS Staining Kit (Roche). Data represent means and SD of at least three independent experiments.

(B) UVC-light-induced growth inhibition of human *ERCC4*-deficient immortal fibroblast cell lines (from XP individual XP2YO and FA individual 1333) transduced with lentiviral particles carrying cDNA coding for wild-type XPF. The results are expressed as a percentage of viable UVC-light-treated cells relative to untreated controls. Data represent means and SD of two independent experiments.

(C) MMC sensitivity of human *ERCC4*-deficient primary fibroblasts from XP, FA, and XFE individuals (XP42RO, 1333, and XP51RO, respectively). Data represent means and SD of two independent experiments.

(D) DEB-induced chromosome-fragility test in human *ERCC4*-deficient primary fibroblasts from XP, FA, and XFE individuals (XP42RO, 1333, and XP51RO, respectively).

(E) MMC-induced G2/M cell-cycle arrest in the same cells as in (D). Experiments presented in (D) and (E) were performed as reported earlier.¹⁰

(Figure 1C). As expected, the truncated XPF was undetectable with an antibody against the C-terminal HhH₂ domain of XPF (amino acids 866–916, data not shown). Interestingly, the truncated XPF was absent in a MMC-resistant lymphoblastoid cell line (1333R) generated by long-term exposure to MMC, and near-normal XPF levels were detected in this reverted cell line (Figure S1C). PCR amplification and sequence analysis revealed that the 28 bp duplication had disappeared in 1333R (Figure S1D) and had thus restored the wild-type sequence. Both the inherited duplication and the somatic reversion might have been triggered by an inverted 5 bp repeat flanking the region.

Genetic complementation of MMC sensitivity in lymphoblasts from both FA individuals was achieved by lentiviral transduction of wild-type *ERCC4* cDNA (Figures 1D and 1E). In addition, we expressed wild-type and mutant human *ERCC4* cDNAs in embryonic fibroblasts (MEFs) from *Ercc4* (*Xpf*)-null mice. We found that ectopic expression of *ERCC4* mutants encoding p.Leu230Pro and p.Arg689Ser did not complement MMC sensitivity of these MEFs (Figure 1F), providing additional evidence that the

ERCC4 missense mutations found in both FA individuals inactivate XPF. The genetic and functional data show that mutations in *ERCC4* cause FA in two unrelated nonconsanguineous individuals. Because mutations in *ERCC4* cause an additional FA subtype (FA-Q), we propose *FANCQ* as an alias for *ERCC4*.

ERCC4 mutations have been linked to the skin-photosensitive and nucleotide excision repair (NER)-deficient disorders xeroderma pigmentosum (XP [MIM 278700, 610651, 278720, 278730, 278740, 278760, 278780, and 278750])⁸ and XFE progeroid syndrome (MIM 610965),⁹ and we therefore tried to understand why the identified *ERCC4* variants specifically lead to FA. We hypothesized that these mutants cause an FA phenotype because of a strong deficiency in ICL repair but have sufficient NER activity to prevent clinically relevant skin photosensitivity and other NER-related features. Compared to an XP complementation group C (XP-C) lymphoblast line, FA104 lymphoblasts were indeed not sensitive to UVC light (Figure 2A). Given that UV-light survival experiments are challenging in lymphoblastoid cell lines, we studied skin fibroblasts from individual 1333 (FA104 fibroblasts were not available) and found that the UV-light sensitivity in FA individual 1333 was milder than that in XP complementation group F (XP-F) individual XP2YO (Figure 2B).

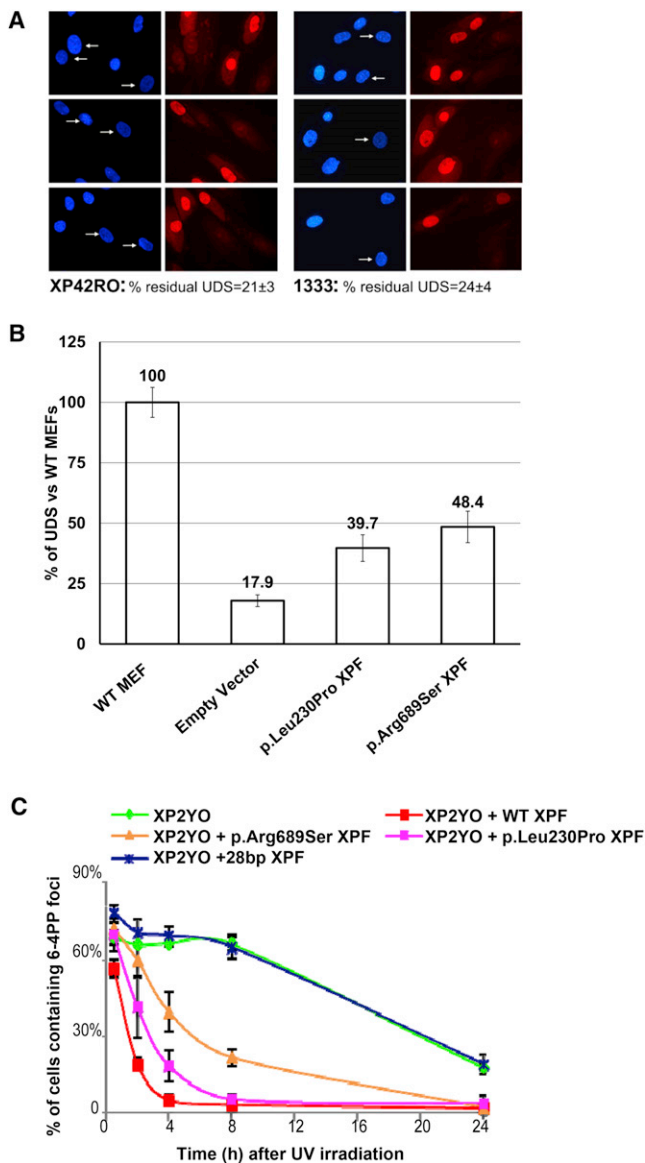


Figure 3. NER Analysis of *ERCC4* Mutants In Vivo

(A and B) In primary fibroblasts, unscheduled DNA synthesis (UDS) representing global NER activity was measured with 5-ethyluracil-thymineless medium (5-TM) as previously described.¹²

(A) XP-F (XP42RO) and FA (1333) cells (arrows) were compared to mixed-in normal fibroblasts preloaded with polystyrene microbeads (no arrows), used as an internal control. UDS signal was quantified from 20–40 random XP-F or FA G1/G2 nuclei and expressed as a percentage of control wild-type cells.

(B) UDS signals in *Ercc4*^{-/-} MEFs measured as in (A) are expressed as a percentage of control wild-type MEFs. *Ercc4*^{-/-} cells were stably expressing an empty vector or one of various *ERCC4* cDNAs (wild-type or encoding p.Leu230Pro or p.Arg689Ser).

(C) Repair kinetics of UV-light-induced DNA damage by FA-specific *ERCC4* mutants in *ERCC4*- and NER-deficient human cells (XP2YO). Cells expressing wild-type XPF, p.Arg689Ser or p.Leu230Pro altered XPF, or XPF resulting from the 28 bp duplication were locally irradiated with UV light, cultured for the indicated times, fixed and stained for 6-4 PPs, and tagged with HA with the use of specific antibodies. Data represent the percentage of cells with 6-4 PP spots at various time points; means and SD of at least two independent experiments are shown. For each experiment, 100 cells were counted.

In addition, the FA-specific XPF alterations p.Leu230Pro and p.Arg689Ser rescued 100% of the UVC sensitivity of XP2YO fibroblasts (Figure S2A) and approximately 80% of the UVC-light sensitivity of *Ercc4*-null MEFs (Figure S2B) but were both unable to complement MMC sensitivity (Figure 1F). Furthermore, XFE and 1333 fibroblasts responded typically like FA cells upon MMC-induced survival (Figure 2C), DEB-induced chromosome breakage (Figure 2D), and MMC-induced G2-phase arrest (Figure 2E), whereas XP-F cells showed milder MMC sensitivity and lacked DEB-induced chromosome fragility and MMC-induced cell-cycle arrest (Figure 2C–2E). Previous experiments in Chinese hamster ovary cells also demonstrated that the XFE-specific p.Arg153Pro altered XPF does not rescue MMC or UV-light sensitivity.¹¹ Therefore, we conclude that XP, XFE, and FA cells with *ERCC4* mutations clearly have a distinct response to UV light and MMC (Table S2).

To further investigate the extent of NER deficiency in the FA-affected individuals, we measured UV-light-induced unscheduled DNA synthesis (UDS) in primary skin fibroblasts from individual 1333 and from an XP-F individual (XP42RO) with mild clinical UV-light sensitivity and found 24 ± 4% and 21 ± 3% residual UDS activity, respectively (Figure 3A). We also determined UDS in *Ercc4*-null MEFs expressing the FA-specific XPF alterations p.Leu230Pro or p.Arg689Ser. The levels of UDS activity were 39.7% and 48.4% of the normal mean for p.Leu230Pro and p.Arg689Ser altered XPF, respectively (Figure 3B), enough to complement 80% of UVC-light sensitivity of these MEFs (Figure S2B). In XPF-deficient human XP2YO fibroblasts, p.Leu230Pro and p.Arg689Ser altered XPF rather efficiently corrected the defective removal of 6-4 photoproducts (PPs) at sites of local UV damage (Figure 3C). In contrast, XP2YO cells expressing the *ERCC4* mutant with the 28 bp duplication were completely deficient in NER activity, as predicted from the disruption of the ERCC1- and DNA-binding domain of this truncated protein. The studies presented in Figures 2 and 3 demonstrate that FA cells with *ERCC4* mutations are fully deficient in ICL repair but retain significant levels of NER activity.

Cell lines from XP-F individuals show a characteristic failure of the altered XPF to properly translocate to the nucleus through aggregation of the protein in the cytoplasm.¹¹ This feature is evident for XP-causing mutations and accentuated in cells from the individual with XFE syndrome. However, FA-causing XPF missense altered proteins can actually translocate to the nucleus, where they are recruited to sites of active NER (Figures S3A and S3B) and can interact with SLX4 and ERCC1 (Figures S3C and S3D). These results might be functionally important, given that a recent article reports that SLX4 interaction with XPF is crucial for ICL repair and that SLX4-knockout mice phenocopy FA.¹³ Using *Xenopus* extracts, J.C. Walter's group reported that the FA upstream pathway genes are required to regulate a nuclease that makes DNA incisions near the ICL.¹⁴ Given that FA-specific altered XPF proteins can

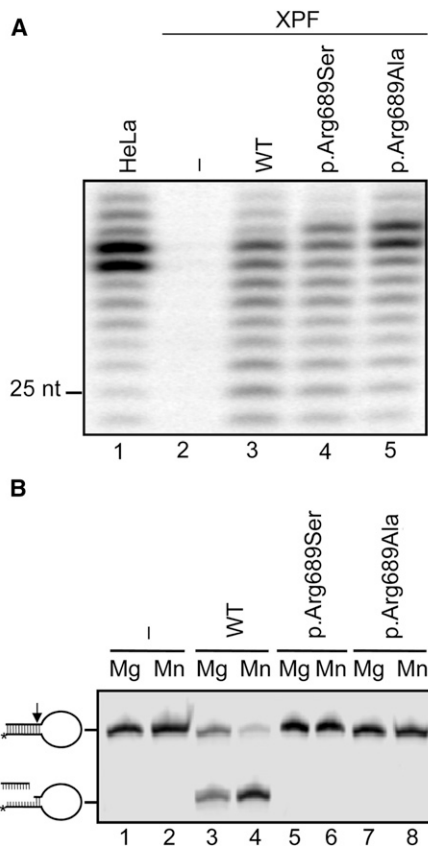


Figure 4. Nuclease Activity of Altered XPF

(A) NER activity of wild-type and altered ERCC1-XPF dimer. A plasmid containing a site-specific 1,3-intrastrand cis-Pt DNA crosslink was incubated with whole-cell extracts from HeLa cells or XPF-deficient cells (XP2YO) complemented with recombinant ERCC1-XPF purified from Sf9 insect cells as reported.¹³ The excised DNA fragments of 24–32 nucleotides are shown. The position of a 25-mer is indicated.

(B) Incision of a stem-loop substrate with wild-type and altered XPF. The 3' Cy5-labeled substrate was incubated with recombinant ERCC1-XPF in the presence of 2 mM MgCl₂ or 0.4 mM MnCl₂, and the products were analyzed by denaturing PAGE. The incision reaction was performed essentially as described earlier.^{5,11}

reach the site of damage, we then investigated their ability to cleave DNA. For this aim, the p.Arg689Ser altered XPF was purified as a heterodimer with ERCC1 as previously described.⁵ Subsequently, NER reactions were performed with the purified altered protein, extracts from XPF-deficient XP2YO cells, and a plasmid containing an NER substrate (1,3-cisplatin intrastrand crosslink).¹⁵ Consistent with the functional data above, the purified heterodimer composed of ERCC1 and p.Arg689Ser XPF is proficient in the excision step of NER similarly to wild-type XPF, given that it restored the ability to cleave and remove the site-specific intrastrand crosslink from the plasmid in XP2YO cell extracts (Figure 4A). Nevertheless, the excision reaction is not perfect given that the excised fragments are, on average, 1 nucleotide longer than expected from a normal reaction with wild-type-XPF-ERCC1 dimer (Figure 4A, lane 4). We also performed in vitro nuclease

activity assays with purified ERCC1-p.Arg689Ser-XPF and ERCC1-p.Arg689Ala-XPF on a stem-loop model DNA substrate. Unlike wild-type XPF and altered XPF proteins causing XP (p.Arg799Trp) or XFE progeroid syndrome (p.Arg153Pro),¹¹ p.Arg689Ser XPF is unable to cleave such a substrate (Figure 4B), indicating that the nuclease-type activity of p.Arg689Ser XPF is grossly abnormal. Unfortunately, we could not perform these biochemical experiments with the p.Leu230Pro altered XPF because we were unable to express and purify ERCC1-p.Leu230Pro-XPF as a result of its low stability and tendency to aggregate. We finally checked whether the FA-specific altered XPF proteins ectopically expressed in *Erc4*-null MEFs can perform the incision step of ICL repair. Both p.Leu230Pro and p.Arg689Ser altered XPF completely restored the incision defect of *Erc4*-null MEFs, as measured by the COMET assay (data not shown), but the cells remained hypersensitive to ICLs (Figure 1F). Although additional biochemical experiments are required, our results suggest that the ICL sensitivity of individuals FA104 and 1333 is not directly linked to the absence of XPF nuclease activity. It seems unlikely that the defect is a downstream step of homologous recombination because FA104 and 1333 cells are not sensitive to PARP inhibitors and are normal in Rad51 focus formation. Given that the nuclease activity of the FA-specific p.Arg689Ser altered XPF is grossly abnormal, it is tempting to speculate that the ICL-unhooking step in these FA cells leaves an intermediate aberrant substrate that is irreparable by subsequent ICL-repair factors.

Our genetic, biochemical, and functional studies, along with the characterization of previous *ERCC4* mutations causing XP-F and XFE, provide a model for the mechanistic understanding of how mutations in *ERCC4* lead to three distinct diseases (Table S2). Most of the presently known XP-F individuals suffer from a relatively mild form of XP.¹⁶ Cells from these individuals have a reduced level of XPF in the nucleus because the altered XPF has a tendency to aggregate in the cytoplasm.¹¹ This reduced level of nuclear XPF is insufficient to mediate complete NER, but it still has enough ICL-repair-specific functions to prevent chromosome fragility, cell-cycle arrest, and subsequent FA clinical manifestations. A second set of *ERCC4* mutations, characterized in this study, allow localization of the protein to the nucleus, where they exert a certain level of NER activity but are fully deficient in ICL repair. p.Arg689Ser XPF is a stable and NER-proficient protein with an active site structure that prevents it from properly processing ICL-repair intermediates. p.Leu230Pro XPF is more similar to the products of previously described *ERCC4* mutations in that it is less stable and might have a tendency to aggregate in the cytoplasm. However, sufficient amounts of the protein are properly folded and reach chromatin, where it appears to have some activity in the removal of 6-4 PPs. Residual NER activity in the skin tissue of individual 1333 in vivo might explain why this individual has no clinically relevant skin photosensitivity,

although we cannot exclude that dermatological problems will arise later in life. A final category of *ERCC4* mutations is associated with XFE progeroid syndrome, which is characterized by very low levels of nuclear XPF, apparently insufficient to support either NER or ICL repair. Importantly, the only XFE-affected individual described suffered from both skin photosensitivity and anemia^{9,16} and shared some cellular features with XP (NER defect and UV-light sensitivity) and FA (extreme ICL sensitivity, DEB-induced chromosome fragility, and MMC-induced cell-cycle arrest), suggesting that XFE syndrome is characterized by a combination of XP and FA manifestations (Table S2). Exhaustion of hematopoietic stem cells is also an attribute of *ERCC1-XPF* hypomorphic mice that mimic XFE (Laura Niedernhofer, personal communication). Microsomy, microcephaly, and liver fibrosis were likewise observed in FA individual 1333, in *Ercc1*- and *Ercc4*-deficient mice, and in the unique *ERCC1*-deficient individual, who all lack ICL-repair functions.^{17–21}

In a broader sense, this study demonstrates that depending on the type of *ERCC4* mutation and the balance between NER and ICL-repair activities, affected individuals present with one of three clinically distinct disorders. This resembles the case of *XPD*, which is involved in XP complementation group D, trichothiodystrophy (MIM 601675), or Cockayne syndrome (MIM 216400) depending on the type of mutation,²² and highlights the value of characterizing rare genetic disorders for gaining insight into the mechanisms of genome maintenance and human disease. XPF has a central role in preventing genome instability, cancer, BMF, developmental abnormalities, and premature aging. Like those of other breast and ovarian cancer susceptibility genes mutated in FA,^{23,24} the product of *ERCC4* also acts downstream of *FANCD2* monoubiquitination. Therefore, it is important to study *FANCD2* as a candidate gene in hereditary breast and ovarian cancer.

Supplemental Data

Supplemental Data include three figures and two tables and can be found with this article online at <http://www.cell.com/AJHG>.

Acknowledgments

The use of *FANCD2* as an alias for *ERCC4* was approved by the HUGO Gene Nomenclature Committee. We would like to thank the families affected by Fanconi anemia and their clinicians for providing samples and clinical data, as well as María A. Blasco (Centro Nacional de Investigaciones Oncológicas, Madrid) for providing *Ercc4*-deficient mouse embryonic fibroblasts. The J.A.B. laboratory is funded by grants from European Program “7FWP, Health” (PERSIST; agreement 222878), the Spanish Ministry of Science and Innovation (Refs110-90.1 and SAF 2009-07164), Programa RETICS-RD06/0010/0015 ISCIII, and Fundación Botín. O.D.S. acknowledges funding from the National Institutes of Health (GM080454 and CA092584). C.S. is funded by CCA/V-ICI Amsterdam. D.S. and B.S. received grants from the Deutsche Fanconi-Anaemie-Hilfe, Aktionskreis Fanconi-Anaemie, and the Schroeder-Kurth Fund. J.S.’s laboratory is funded by the

Generalitat de Catalunya (SGR0489-2009), the ICREA-Academia award, the Spanish Ministry of Science and Innovation (Centre for Biomedical Network Research on Rare Diseases [CIBERER] CB06/07/0023, SAF2009-11936, and SAF2012-31881), and the European Regional Development FEDER Funds. CIBERER is an initiative of the Instituto de Salud Carlos III, Spain.

Received: November 13, 2012

Revised: March 5, 2013

Accepted: April 5, 2013

Published: April 25, 2013

Web Resources

The URLs for data presented herein are as follows:

Online Mendelian Inheritance in Man (OMIM), <http://www.omim.org>

RefSeq, <http://www.ncbi.nlm.nih.gov/RefSeq>

References

1. Stoepker, C., Hain, K., Schuster, B., Hillhorst-Hofstee, Y., Rooimans, M.A., Steltenpool, J., Oostra, A.B., Eirich, K., Korthof, E.T., Nieuwint, A.W., et al. (2011). *SLX4*, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nat. Genet.* **43**, 138–141.
2. Kim, Y., Lach, F.P., Desetty, R., Hanenberg, H., Auerbach, A.D., and Smogorzewska, A. (2011). Mutations of the *SLX4* gene in Fanconi anemia. *Nat. Genet.* **43**, 142–146.
3. Antonio Casado, J., Callén, E., Jacome, A., Río, P., Castella, M., Lobitz, S., Ferro, T., Muñoz, A., Sevilla, J., Cantalejo, A., et al. (2007). A comprehensive strategy for the subtyping of patients with Fanconi anaemia: conclusions from the Spanish Fanconi Anemia Research Network. *J. Med. Genet.* **44**, 241–249.
4. Deans, A.J., and West, S.C. (2009). *FANCD2* connects the genome instability disorders Bloom’s Syndrome and Fanconi Anemia. *Mol. Cell* **36**, 943–953.
5. Enzlin, J.H., and Schärer, O.D. (2002). The active site of the DNA repair endonuclease XPF-*ERCC1* forms a highly conserved nuclease motif. *EMBO J.* **21**, 2045–2053.
6. Almarza, E., Río, P., Meza, N.W., Aldea, M., Agirre, X., Guenechea, G., Segovia, J.C., and Bueren, J.A. (2007). Characteristics of lentiviral vectors harboring the proximal promoter of the *vav* proto-oncogene: a weak and efficient promoter for gene therapy. *Mol. Ther.* **15**, 1487–1494.
7. de Laat, W.L., Sijbers, A.M., Odijk, H., Jaspers, N.G., and Hoeijmakers, J.H. (1998). Mapping of interaction domains between human repair proteins *ERCC1* and XPF. *Nucleic Acids Res.* **26**, 4146–4152.
8. Sijbers, A.M., de Laat, W.L., Ariza, R.R., Biggerstaff, M., Wei, Y.F., Moggs, J.G., Carter, K.C., Shell, B.K., Evans, E., de Jong, M.C., et al. (1996). Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86**, 811–822.
9. Niedernhofer, L.J., Garinis, G.A., Raams, A., Lalai, A.S., Robinson, A.R., Appeldoorn, E., Odijk, H., Oostendorp, R., Ahmad, A., van Leeuwen, W., et al. (2006). A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* **444**, 1038–1043.
10. Trujillo, J.P., Mina, L.B., Pujol, R., Bogliolo, M., Andrieux, J., Holder, M., Schuster, B., Schindler, D., and Surrallés, J.

- (2012). On the role of FAN1 in Fanconi anemia. *Blood* 120, 86–89.
11. Ahmad, A., Enzlin, J.H., Bhagwat, N.R., Wijgers, N., Raams, A., Appeldoorn, E., Theil, A.F., J Hoeijmakers, J.H., Vermeulen, W., J Jaspers, N.G., et al. (2010). Mislocalization of XPF-ERCC1 nuclease contributes to reduced DNA repair in XP-F patients. *PLoS Genet.* 6, e1000871.
 12. Limsirichaikul, S., Niimi, A., Fawcett, H., Lehmann, A., Yamashita, S., and Ogi, T. (2009). A rapid non-radioactive technique for measurement of repair synthesis in primary human fibroblasts by incorporation of ethynyl deoxyuridine (EdU). *Nucleic Acids Res.* 37, e31.
 13. Crossan, G.P., van der Weyden, L., Rosado, I.V., Langevin, F., Gaillard, P.H., McIntyre, R.E., Gallagher, F., Kettunen, M.I., Lewis, D.Y., Brindle, K., et al.; Sanger Mouse Genetics Project. (2011). Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. *Nat. Genet.* 43, 147–152.
 14. Knipscheer, P., Räschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Schärer, O.D., Elledge, S.J., and Walter, J.C. (2009). The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* 326, 1698–1701.
 15. Moggs, J.G., Yarema, K.J., Essigmann, J.M., and Wood, R.D. (1996). Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *J. Biol. Chem.* 271, 7177–7186.
 16. Gregg, S.Q., Robinson, A.R., and Niedernhofer, L.J. (2011). Physiological consequences of defects in ERCC1-XPF DNA repair endonuclease. *DNA Repair (Amst.)* 10, 781–791.
 17. Weeda, G., Donker, I., de Wit, J., Morreau, H., Janssens, R., Vissers, C.J., Nigg, A., van Steeg, H., Bootsma, D., and Hoeijmakers, J.H. (1997). Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr. Biol.* 7, 427–439.
 18. McWhir, J., Selfridge, J., Harrison, D.J., Squires, S., and Melton, D.W. (1993). Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat. Genet.* 5, 217–224.
 19. Tian, M., Shinkura, R., Shinkura, N., and Alt, F.W. (2004). Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. *Mol. Cell. Biol.* 24, 1200–1205.
 20. Jaspers, N.G., Raams, A., Silengo, M.C., Wijgers, N., Niedernhofer, L.J., Robinson, A.R., Giglia-Mari, G., Hoogstraten, D., Kleijer, W.J., Hoeijmakers, J.H., and Vermeulen, W. (2007). First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. *Am. J. Hum. Genet.* 80, 457–466.
 21. Gregg, S.Q., Gutiérrez, V., Robinson, A.R., Woodell, T., Nakao, A., Ross, M.A., Michalopoulos, G.K., Rigatti, L., Rothmel, C.E., Kamileri, I., et al. (2012). A mouse model of accelerated liver aging caused by a defect in DNA repair. *Hepatology* 55, 609–621.
 22. Cleaver, J.E., Lam, E.T., and Revet, I. (2009). Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat. Rev. Genet.* 10, 756–768.
 23. Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D., Freund, M., Lichtner, P., Hartmann, L., Schaal, H., et al. (2010). Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat. Genet.* 42, 410–414.
 24. Levy-Lahad, E. (2010). Fanconi anemia and breast cancer susceptibility meet again. *Nat. Genet.* 42, 368–369.

Anexo 3

Osorio A, **Bogliolo M**, Fernández V, Barroso A, de la Hoya M, Caldés T, Lasa A, Ramón y Cajal T, Santamariña M, Vega A, Quiles F, Lázaro C, Díez O, Fernández D, González-Sarmiento R, Durán M, Piqueras JF, Marín M, Pujol R, Surrallés J, Benítez J. **Evaluation of rare variants in the new fanconi anemia gene ERCC4(FANCC) as familial breast/ovarian cancer susceptibility alleles.** Hum Mutat. 2013Dec;34(12):1615-8.

Evaluation of Rare Variants in the New Fanconi Anemia Gene *ERCC4* (*FANCO*) as Familial Breast/Ovarian Cancer Susceptibility Alleles

Ana Osorio,^{1,2†} Massimo Bogliolo,^{2,3†} Victoria Fernández,¹ Alicia Barroso,¹ Miguel de la Hoya,⁴ Trinidad Caldes,⁴ Adriana Lasa,⁵ Teresa Ramón y Cajal,⁶ Marta Santamaría,^{2,7} Ana Vega,^{2,8} Francisco Quiles,⁹ Conxi Lázaro,⁹ Orland Díez,¹⁰ Daniel Fernández,¹¹ Rogelio González-Sarmiento,¹¹ Mercedes Durán,¹² José Fernández Piqueras,^{2,13} María Marín,³ Roser Pujol,^{2,3} Jordi Surrallés,^{2,3*} and Javier Benítez^{1,2**}

¹Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre, CNIO, Madrid, Spain; ²Centre for Biomedical Network Research on Rare Diseases (CIBERER), Spain; ³Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain; ⁴Molecular Oncology Laboratory, Hospital Clínico San Carlos. Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Barcelona, Spain; ⁵Genetics Service, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; ⁶Oncology Service, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; ⁷Grupo de Medicina Xenómica -USC, University of Santiago de Compostela, CIBERER, IDIS, Santiago de Compostela, Spain; ⁸Fundación Pública Galega de Medicina Xenómica-SERGAS Santiago de Compostela, Spain; ⁹Hereditary Cancer Program, Catalan Institute of Oncology (ICO), Hospital Duran i Reynals, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet, Barcelona, Spain; ¹⁰Oncogenetics Laboratory, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, and University Hospital of Vall d'Hebron, Barcelona, Spain; ¹¹Instituto de Biología Molecular y Celular del Cáncer (IBMCC), laboratory 14, Universidad de Salamanca-CSIC, Salamanca, Spain; ¹²Institute of Biology and Molecular Genetics, Universidad de Valladolid (IBGM-UVA), Valladolid, Spain; ¹³Centro de Biología Molecular Severo Ochoa (CBM), Consejo Superior de Investigaciones Científicas-Universidad Autónoma de Madrid, Spain

Communicated by Christopher G. Mathew

Received 31 May 2013; accepted revised manuscript 28 August 2013.

Published online 11 September 2013 in Wiley Online Library (www.wiley.com/humanmutation). DOI: 10.1002/humu.22438

ABSTRACT: Recently, it has been reported that biallelic mutations in the *ERCC4* (*FANCO*) gene cause Fanconi anemia (FA) subtype FA-Q. To investigate the possible role of *ERCC4* in breast and ovarian cancer susceptibility, as occurs with other FA genes, we screened the 11 coding exons and exon–intron boundaries of *ERCC4* in 1573 index cases from high-risk Spanish familial breast and ovarian cancer pedigrees that had been tested negative for *BRCA1* and *BRCA2* mutations and 854 controls.

The frequency of *ERCC4* mutation carriers does not differ between cases and controls, suggesting that *ERCC4* is not a cancer susceptibility gene. Interestingly, the prevalence of *ERCC4* mutation carriers (one in 288) is similar to that reported for *FANCA*, whereas there are approximately 100-fold more FA-A than FA-Q patients, indicating that most biallelic combinations of *ERCC4* mutations are embryo lethal. Finally, we identified additional bone-fide FA *ERCC4* mutations specifically disrupting interstrand cross-link repair.

Hum Mutat 34:1615–1618, 2013. © 2013 Wiley Periodicals, Inc.

KEY WORDS: Fanconi anemia; breast cancer; *ERCC4*; *FANCO*; XPF

†These authors contributed equally to this work.

*Correspondence to: Jordi Surrallés, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. E-mail: jordi.surralles@uab.es

**Correspondence to: Javier Benítez, Human Genetics Programme Spanish National Cancer Centre (CNIO) C/Melchor Fernández Almagro 3, 28029 Madrid, Spain. E-mail: jbenitez@cnio.es

Contract grant sponsors: Mutua Madrileña Foundation (FMMA); Intrasalud PI12/00070; Acción Cooperativa y Colaborativa Intramural-CIBER2012; SAF2010-20493; Generalitat de Catalunya (SGR0489-2009), the ICREA-Academia award, the Spanish Ministry of Economy and Competitiveness (CIBERER CB06/07/0023, SAF2009-11936 and SAF2012-31881); the European Regional Development FEDER Funds; Fondo de Investigación Sanitaria (FIS) (PI 12/00539); European Regional Development FEDER Funds; Asociación Española Contra el Cáncer; Spanish Health Research Fund; Carlos III Health Institute; Catalan Health Institute and Autonomous Government of Catalonia (ISCIII/RETIC RD06/0020/1051, RD12/0036/008, PI10/01422, PI10/00748, and 2009SGR290). Acción Cooperativa y Colaborativa Intramural-CIBERER 2012, the OncoCycle Programme (S2011/BMD-2470) and SAF2012-36556; and Fondo de Investigación Sanitaria (PI10/00219)

The *ERCC4* (*FANCO*) gene (MIM #133520) encodes for a DNA repair endonuclease (XPF) that plays essential roles in nucleotide excision repair (NER) and interstrand cross-link repair (ICLR) [Gregg et al., 2011]. Biallelic mutations in *ERCC4* have been linked to Xeroderma Pigmentosum (XP; MIMs #278700, #610651, #278720, #278730, #278740, #278760, #278780, and #278750) [Sijbers et al., 1996] and progeria (XFE; MIM #610965) [Niedernhofer et al., 2006] diseases. Very recently, it has been published that biallelic mutations in the gene are also causative of disease in previously unclassified Fanconi anemia (FA) patients subtype FA-Q and, therefore, the use of *FANCO* as an alias for *ERCC4* was approved by the HUGO Gene Nomenclature Committee [Bogliolo et al., 2013]. Individuals harboring mutations in *ERCC4* show high variability on their clinical manifestations, ranging from mild XP symptoms (sun sensitivity, freckling of the skin, and basal or squamous cell carcinomas) to the dramatic accelerated aging symptoms of a progeroid

Table 1. Unique Genetic Variants Identified in the *ERCC4* Gene in 1573 Cases and 854 Controls

Exon	Nucleotide change ^a	Protein change	Previously described	MAF reported ^c	Phenotype	Predicted to affect protein function ^e	Functional assay ^f
3	c.448C>T ^b	p.Arg150Cys	rs145402255	0.0	Case	Yes	Mild effect on ICLR
3	c.540_541delAG	p.Arg180SerfsX30	–	–	Control ^d	Yes	–
IVS3	c.584+1G>A ^b	p.Ile131SerfsX2	–	–	Case	Yes	–
5	c.800G>A	p.Arg267His	rs143479220	0.0001	Case	No	–
9	c.1861A>G	p.Thr621Ala	–	–	Case	No	–
11	c.2065C>A	p.Arg689Ser	rs149364215	0.0001	Control	Yes	Disrupts ICLR
11	c.2291delG	p.Ser764ThrfsX53	–	–	Control	Yes	–
11	c.2357C>T ^b	p.Ser786Phe	rs143081574	–	Case	Yes	Disrupts ICLR

^aMutation nomenclature listed uses GenBank reference sequences NM_005236.2 with numbering starting at the A of the first ATG, following the HGVS guidelines (www.hgvs.org/mutnomen).

^bApart from the 854 controls in which *XPF/ERCC4* was fully sequenced, these variants were specifically analyzed in 300 additional controls. The mutation gave rise to skipping of exon 3 (r.389_584del) that was confirmed at the cDNA level (data not shown).

^cMAF (minor allele frequency) in European American Population as described in the NHLBI Exome Sequencing Project (ESP) (<http://evs.gs.washington.edu/EVS/>).

^dMutation found in two independent controls.

^eAll mutations giving rise to a PTC were considered as predicted to affect protein function. Missense mutations were previously evaluated using the program CONDEL that predicts the pathogenicity of nonsynonymous variants using a consensus deleteriousness score that combines various tools such as SIFT, Polyphen2, and MutationAssessor (<http://bg.upf.edu/condel/home>). Those missense variants predicted to be deleterious by condel were considered as predicted to affect protein function and this was confirmed by functional analysis; those predicted by condel to be neutral were considered as such and not further evaluated.

^fDetailed in the text and in Figure 1.

syndrome, and finally, as recently shown [Bogliolo et al., 2013] in FA patients, to malformations and progressive bone marrow failure. It has also been demonstrated that a balance between NER and ICLR activities due to different effects caused by distinct mutations in *ERCC4* determined the final clinical outcome [Bogliolo et al., 2013]. On the other hand, given that FA-Q patients are normal in FANCD2 monoubiquitination, it has been suggested that monoallelic mutations in *ERCC4* could be linked to the breast and ovarian hereditary cancer syndrome as it occurs with other FA genes such as *BRCA2/FANCD1* [Howlett et al., 2002], *BRIP1/FANCF* [Seal et al., 2006], *PALB2/FANCF* [Rahman et al., 2007], and *RAD51C/FANCO* [Meindl et al., 2010; Osorio et al., 2012] (MIMs #300514, #610832, #614082, and #614083, respectively) [Levy-Lahad, 2010].

To investigate the possible role of *ERCC4* in breast and ovarian cancer susceptibility, we screened, by DHPLC (denaturing high-performance liquid chromatography) and direct sequencing, the 11 coding exons and exon–intron boundaries of the *ERCC4* gene in 1573 index cases from high-risk Spanish familial breast and ovarian cancer pedigrees that had been tested negative for mutations in *BRCA1* and *BRCA2* and 854 controls without personal or familial antecedents of cancer. Criteria for inclusion of cases and controls, and methods of screening for mutations in *BRCA1/2* have been previously published [Bonache et al., 2013; Fernandez-Rodriguez et al., 2012; Osorio et al., 2012; Romero et al., 2011]. We identified five and four unique variants among cases and controls among which three and four, respectively, were considered as putatively deleterious. Missense mutations were evaluated using the program CONDEL that predicts the pathogenicity of nonsynonymous variants using a consensus deleteriousness score that combines various tools such as SIFT, Polyphen2, and MutationAssessor (<http://bg.upf.edu/condel/home>). Those missense variants predicted to be deleterious by CONDEL were considered as predicted to affect protein function and this was studied later by functional analysis; those predicted to be neutral were considered as such and not further evaluated (Table 1). Regarding cases, one of the mutations found, c.584+1G>A in intron 3, was confirmed to cause skipping of exon 3 of the gene and a premature stop codon (PTC) (data not shown). The other two mutations, p.Arg150Cys in exon 3 and p.Ser786Phe in exon 11, were later functionally investigated. Regarding controls, we identified two different frameshift mutations c.540_541delAG in exon 3 (found in two independent controls) and c.2291delG in exon 11, both predicted to cause PTCs. Even though

this later PTC is very C-terminal and could potentially result in a shorter but partially functional protein, this mutation results in a truncated XPF protein that lacks the double helix–hairpin–helix (HhH2) domain involved in heterodimerization with ERCC1 and DNA binding [de Laat et al., 1998], very similar to a pathogenic *ERCC4* mutation (c.2371_2398dup28; p.Ile800Thrfs*24) that functionally disrupts NER and ICLR activities [Bogliolo et al., 2013]. The last deleterious variant found in controls was the missense p.Arg689Ser in exon 11 previously found in a FA patient, and demonstrated to cause abnormal nuclease activity and to specifically disrupt ICLR [Bogliolo et al., 2013]. No differences were found regarding the localization of mutations in the gene among cases and controls (Table 1). All variants reported have been submitted to the Leiden Open Variation Database (LOVD).

To evaluate the functional impact of the missense variants, we cloned a HA-tagged wild type (WT) *ERCC4* cDNA in a pBABE-puro retroviral vector (Addgene plasmid 14430, kindly shared by Dr. L.M. Martins) in internal ribosome entry site (IRES) with the GFP cDNA, and the c.448C>T, c.2065C>A, and c.2357C>T variants were introduced by site directed mutagenesis [Bogliolo et al., 2013]. The resulting constructs were transduced in NER and ICLR deficient *Ercc4* KO mouse embryonic fibroblasts (MEFs) and, after puromycin selection, the green cells were sorted to achieve a purity of over 98% by fluorescence activated cell sorting [Bogliolo et al., 2013]. Due to the bicistronic nature of the IRES construct, we were able to assess the stability of the mutant XPF proteins using GFP as a reference, since both proteins are encoded by the same mRNA [Pelletier and Sonenberg, 1988]. The p.Arg689Ser variant reduced by 40% the stability of XPF, whereas the p.Arg150Cys and p.Ser786Phe variants had no impact on protein stability (Fig. 1A and B). Ultraviolet radiation subtype C (UVC) sensitivity of Xpf KO MEFs was complemented with the expression of both p.Arg150Cys and p.Ser786Phe-XPF (Fig. 1C), and only the Xpf KO MEFs expressing p.Ser786Phe-XPF or p.Arg689Ser showed a FA phenotype in terms of mitomycin C (MMC) sensitivity (Fig. 1D), MMC-induced cell cycle arrest at the G2/M phase (Fig. 1E), and DEB-induced chromosome fragility (Fig. 1F). These data confirm that, resembling p.Arg689Ser, p.Ser786Phe specifically disrupts ICLR and, therefore, is a bone-fide FA mutation. Interestingly, both mutations are located in the nuclease domain of XPF. Despite a mild MMC sensitivity (Fig. 1D and E), *Ercc4* KO MEFs expressing p.Arg150Cys-XPF did not show DEB-induced chromosome fragility (Fig. 1F). These data,

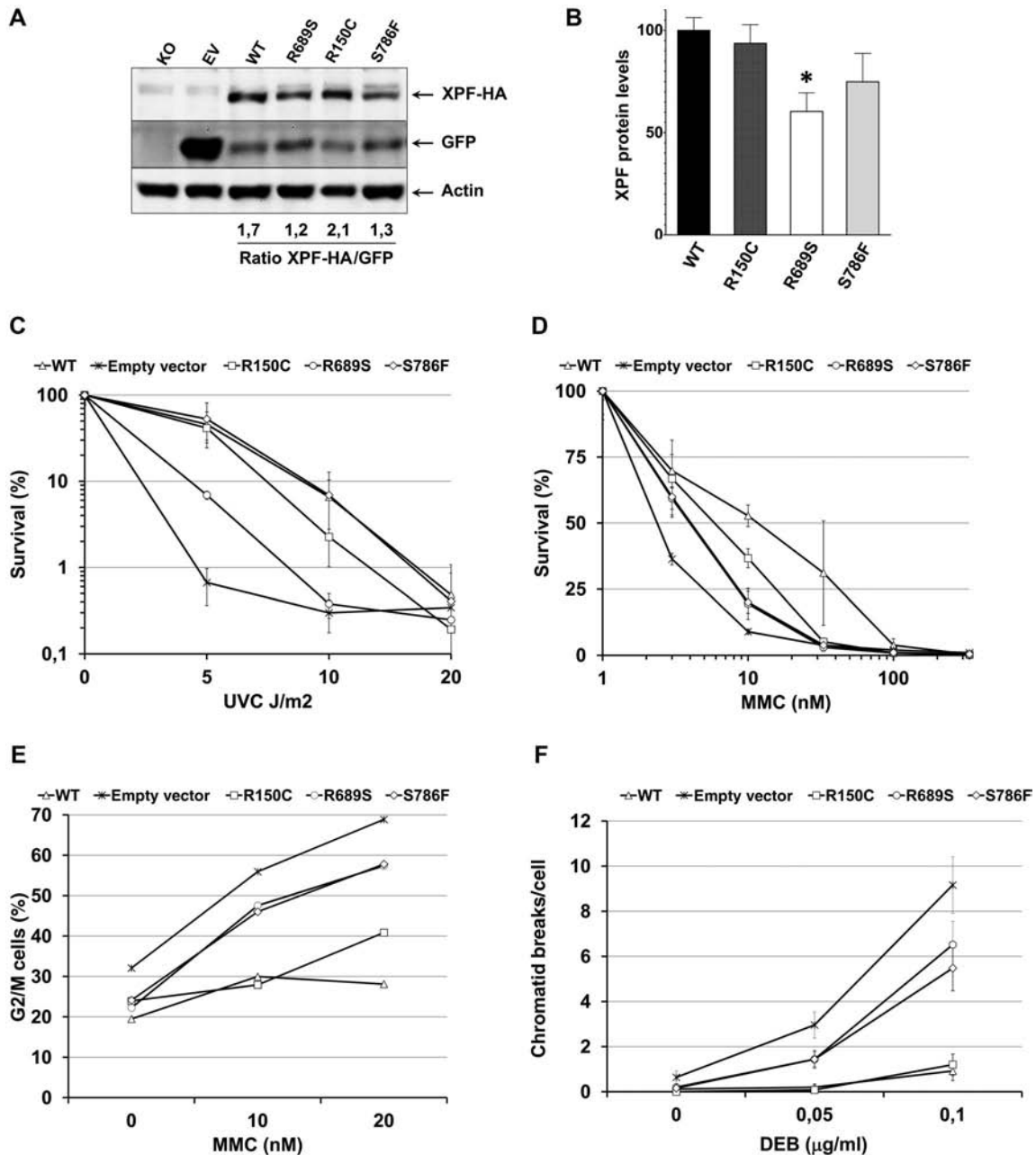


Figure 1. Functional studies of *ERCC4* missense variants. **A:** Western blot analysis showing HA-XPF and GFP expressions in the transduced MEFs as a measure of XPF protein stability. XPF levels were quantified with ImageJ software and are expressed as a ratio relative to the GFP levels. **B:** XPF levels were quantified as in (A) and expressed as a percentage of the WT. Mean and SEM of six experiments are shown (* $p = 0.005$ student's t -test). **C:** UVC sensitivity of *Ercc4* KO MEFs expressing the different alleles of *ERCC4*. Cells were analyzed after 72 hr of irradiation. Data represent means and SD of three independent experiments. **D:** MMC sensitivity in the same cells as in (C). Data represent means and SD of two independent experiments. **E:** MMC-induced G2/M cell cycle arrest in the same cells as in (C). **F:** DEB-induced chromosome fragility test arrest in the same cells as in (C). Chromatid breaks were scored in 25 metaphase per point. Methods for western blot, site-directed mutagenesis, cDNA transduction, antibodies used, and the experiments presented in (D), (E), and (F) were performed as reported earlier [Bogliolo et al., 2013; Trujillo et al., 2012].

together with the protein stability and UVC sensitivity data, indicate a null impact of the c.448C>T variant on XPF NER functions and a mild effect on ICLR activity.

In conclusion, the frequency of Spanish individuals heterozygous for pathogenic mutations in the *ERCC4* gene is approximately 0.3%, and it does not differ between familial breast/ovarian cancer patients and healthy controls ($p = 0.251$), suggesting that monoallelic muta-

tions in *ERCC4* are not linked to cancer susceptibility in the general population. Similar results were found with *SLX4* that, like *ERCC4*, acts downstream FANCD2 monoubiquitination but upstream the homologous recombination step of ICLR [Fernandez-Rodriguez et al., 2012]. The prevalence of *ERCC4* mutation carriers (one in 288) is similar to that reported for *FANCA*. However, there are approximately 100-fold more FA-A than FA-Q individuals, suggesting

that over 90% of biallelic combinations of *ERCC4* mutations are embryo lethal in humans. All reported XP patients subtype XPF worldwide have at least one missense mutation disrupting NER, whereas all missense mutations found in 2.427 Spanish individuals have substantial NER activity explaining why there are no reported XPF families in Spain.

Acknowledgments

This study was approved by the Institutional Committee on Ethical Research in Human Samples and proper informed consent was obtained. CIBERER is an initiative of the Instituto de Salud Carlos III, Spain. We wish to thank all the members of the ICO Hereditary Cancer Program team led by Dr. Gabriel Capella.

All authors of this manuscript made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data, collaborated in drafting the article or revising it critically for important intellectual content and gave final approval of the version to be published.

Disclosure statement: The authors declare no conflict of interest.

References

- Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A, Trujillo JP, Minguillon J, Ramirez MJ, Pujol R, Casado JA, Banos R, et al. 2013. Mutations in *ERCC4*, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet* 92:800–806.
- Bonache S, de la Hoya M, Gutierrez-Enriquez S, Tenes A, Masas M, Balmana J, Diez O. 2013. Mutation analysis of the *SHFM1* gene in breast/ovarian cancer families. *J Cancer Res Clin Oncol* 139:529–532.
- de Laat WL, Sijbers AM, Odijk H, Jaspers NG, Hoeijmakers JH. 1998. Mapping of interaction domains between human repair proteins *ERCC1* and XPF. *Nucleic Acids Res* 26:4146–4152.
- Fernandez-Rodriguez J, Quiles F, Blanco I, Teule A, Feliubadalo L, Valle JD, Salinas M, Izquierdo A, Darder E, Schindler D, Capella G, Brunet J, et al. 2012. Analysis of *SLX4/FANCP* in non-*BRCA1/2*-mutated breast cancer families. *BMC Cancer* 12:84.
- Gregg SQ, Robinson AR, Niedernhofer LJ. 2011. Physiological consequences of defects in *ERCC1-XPF* DNA repair endonuclease. *DNA Repair (Amst)* 10:781–791.
- Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C, Persky N, Grompe M, Joenje H, Pals G, Ikeda H, Fox EA, et al. 2002. Biallelic inactivation of *BRCA2* in Fanconi anemia. *Science* 297:606–609.
- Levy-Lahad E. 2010. Fanconi anemia and breast cancer susceptibility meet again. *Nat Genet* 42:368–369.
- Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, Freund M, Lichtner P, Hartmann L, Schaal H, Ramser J, Honisch E, et al. 2010. Germline mutations in breast and ovarian cancer pedigrees establish *RAD51C* as a human cancer susceptibility gene. *Nat Genet* 42:410–414.
- Niedernhofer LJ, Garinis GA, Raams A, Lalai AS, Robinson AR, Appeldoorn E, Odijk H, Oostendorp R, Ahmad A, van Leeuwen W, Theil AF, Vermeulen W, et al. 2006. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* 444:1038–1043.
- Osorio A, Endt D, Fernandez F, Eirich K, de la Hoya M, Schmutzler R, Caldes T, Meindl A, Schindler D, Benitez J. 2012. Predominance of pathogenic missense variants in the *RAD51C* gene occurring in breast and ovarian cancer families. *Hum Mol Genet* 21:2889–2898.
- Pelletier J, Sonenberg N. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334:320–325.
- Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T, Jayatilake H, McGuffog L, et al. 2007. *PALB2*, which encodes a *BRCA2*-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 39:165–167.
- Romero A, Perez-Segura P, Tosar A, Garcia-Saenz JA, Diaz-Rubio E, Caldes T, de la Hoya M. 2011. A HRM-based screening method detects *RAD51C* germ-line deleterious mutations in Spanish breast and ovarian cancer families. *Breast Cancer Res Treat* 129:939–946.
- Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K, North B, McGuffog L, et al. 2006. Truncating mutations in the Fanconi anemia J gene *BRIP1* are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 38:1239–1241.
- Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, Carter KC, Shell BK, Evans E, de Jong MC, Rademakers S, de Rooij J, et al. 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* 86:811–822.
- Trujillo JP, Mina LB, Pujol R, Bogliolo M, Andrieux J, Holder M, Schuster B, Schindler D, Surrallés J. 2012. On the role of *FAN1* in Fanconi anemia. *Blood* 120: 86–89.

Anexo 4

Peterlongo P, Catucci I, Colombo M, Caleca L, Mucaki E, **Bogliolo M**, Marin M, Damiola F, Bernard L, Pensotti V, Volorio S, Dall'Olio V, Meindl A, Bartram C, Sutter C, Surowy H, Sornin V, Dondon MG, Eon-Marchais S, Stoppa-Lyonnet D, Andrieu N, Sinilnikova OM; GENESIS, Mitchell G, James PA, Thompson E; kConFab; SWE-BRCA, Marchetti M, Verzeroli C, Tartari C, Capone GL, Putignano AL, Genuardi M, Medici V, Marchi I, Federico M, Tognazzo S, Matricardi L, Agata S, Dolcetti R, Puppa LD, Cini G, Gismondi V, Viassolo V, Perfumo C, Mencarelli MA, Baldassarri M, Peissel B, Roversi G, Silvestri V, Rizzolo P, Spina F, Vivinet C, Tibiletti MG, Caligo MA, Gambino G, Tommasi S, Pilato B, Tondini C, Corna C, Bonanni B, Barile M, Osorio A, Benitez J, Balestrino L, Ottini L, Manoukian S, Pierotti MA, Renieri A, Varesco L, Couch FJ, Wang X, Devilee P, Hilbers FS, van Asperen CJ, Viel A, Montagna M, Cortesi L, Diez O, Balmaña J, Hauke J, Schmutzler RK, Papi L, Pujana MA, Lázaro C, Falanga A, Offit K, Vijai J, Campbell I, Burwinkel B, Kvist A, Ehrencrona H, Mazoyer S, Pizzamiglio S, Verderio P, Surralles J, Rogan PK, Radice P. **FANCM c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity and is a familial breast cancer risk factor.** Hum Mol Genet. 2015 Sep 15;24(18):5345-5355.

ASSOCIATION STUDIES ARTICLE

FANCM c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity and is a familial breast cancer risk factor

Paolo Peterlongo^{1,2,*}, Irene Catucci^{1,2}, Mara Colombo², Laura Galeca², Eliseos Mucaki⁶, Massimo Bogliolo⁷, Maria Marin⁷, Francesca Damiola⁸, Loris Bernard^{9,11}, Valeria Pensotti^{1,11}, Sara Volorio^{1,11}, Valentina Dall'Olio^{1,11}, Alfons Meindl¹², Claus Bartram¹³, Christian Sutter¹³, Harald Surowy^{14,15}, Valérie Sornin⁸, Marie-Gabrielle Dondon^{16,17,18}, Séverine Eon-Marchais^{16,17,18}, Dominique Stoppa-Lyonnet^{19,20,21}, Nadine Andrieu^{16,17,18}, Olga M. Sinilnikova^{8,22}, GENESIS[†], Gillian Mitchell^{23,26}, Paul A. James^{23,26}, Ella Thompson^{24,26}, kConFab²⁵, SWE-BRCA[‡], Marina Marchetti²⁸, Cristina Verzeroli²⁵, Carmen Tartari²⁸, Gabriele Lorenzo Capone^{30,31}, Anna Laura Putignano^{30,31}, Maurizio Genuardi^{30,31,32}, Veronica Medici³³, Isabella Marchi³³, Massimo Federico³³, Silvia Tognazzo³⁴, Laura Matricardi³⁴, Simona Agata³⁴, Riccardo Dolcetti³⁵, Lara Della Puppa³⁶, Giulia Cini³⁶, Viviana Gismondi³⁷, Valeria Viassolo³⁷, Chiara Perfumo³⁷, Maria Antonietta Mencarelli^{38,39}, Margherita Baldassarri^{38,39}, Bernard Peissel³, Gaia Roversi³, Valentina Silvestri⁴⁰, Piera Rizzolo⁴⁰, Francesca Spina⁴¹, Caterina Vivinet⁴¹, Maria Grazia Tibiletti⁴², Maria Adelaide Caligo⁴³, Gaetana Gambino⁴³, Stefania Tommasi⁴⁴, Brunella Pilato⁴⁴, Carlo Tondini²⁹, Chiara Corna²⁹, Bernardo Bonanni¹⁰, Monica Barile¹⁰, Ana Osorio^{45,46}, Javier Benitez^{45,46}, Luisa Balestrino⁴¹, Laura Ottini⁴⁰, Siranoush Manoukian³, Marco A. Pierotti⁴, Alessandra Renieri^{38,39}, Liliana Varesco³⁷, Fergus J. Couch⁴⁷, Xianshu Wang⁴⁷, Peter Devilee^{48,49}, Florentine S. Hilbers⁴⁸, Christi J. van Asperen⁵⁰, Alessandra Viel³⁶, Marco Montagna³⁴, Laura Cortesi³³, Orland Diez^{51,52,53}, Judith Balmaña^{52,54}, Jan Hauke⁵⁵, Rita K. Schmutzler⁵⁵, Laura Papi³⁰, Miguel Angel Pujana⁵⁶, Conxi Lázaro⁵⁶,

[†]A full list of collaborators is provided in the Supplementary Material, Note.

[‡]The Swedish BRCA1 and BRCA2 study collaborators; a full list of collaborators is provided in the Supplementary Material, Note.

Received: December 1, 2014. Revised: June 9, 2015. Accepted: June 25, 2015

© The Author 2015. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

Anna Falanga²⁸, Kenneth Offit^{57,58,59}, Joseph Vijai^{57,58,59}, Ian Campbell^{24,26,27}, Barbara Burwinkel^{14,15}, Anders Kvist⁶⁰, Hans Ehrencrona^{61,62}, Sylvie Mazoyer⁸, Sara Pizzamiglio⁵, Paolo Verderio⁵, Jordi Surrallés⁷, Peter K. Rogan⁶ and Paolo Radice^{1,2}

¹IFOM, the FIRC Institute of Molecular Oncology, Milan, Italy, ²Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, ³Unit of Medical Genetics, Department of Preventive and Predictive Medicine, ⁴Scientific Directorate and ⁵Unit of Medical Statistics, Biometry and Bioinformatics, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy, ⁶Department of Biochemistry, Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada, ⁷Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona and Center for Biomedical Network Research on Rare Diseases (CIBERER), Barcelona, Spain, ⁸Cancer Research Centre of Lyon, CNRS UMR5286, INSERM U1052, Université Claude Bernard Lyon 1, Centre Léon Bérard, Lyon, France, ⁹Department of Experimental Oncology and ¹⁰Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia, Milan, Italy, ¹¹Cogentech, Cancer Genetic Test Laboratory, Milan, Italy, ¹²Division of Gynaecology and Obstetrics, Technische Universität München, Munich, Germany, ¹³Institute of Human Genetics, University of Heidelberg, Heidelberg, Germany, ¹⁴Molecular Biology of Breast Cancer, Department of Obstetrics and Gynecology, University Hospital Heidelberg, Heidelberg, Germany, ¹⁵Molecular Epidemiology Group, C080, German Cancer Research Center (DKFZ), Heidelberg, Germany, ¹⁶INSERM, U900, Paris, France, ¹⁷Institut Curie, Paris, France, ¹⁸Mines ParisTech, Fontainebleau, France, ¹⁹Service de Génétique Oncologique, Institut Curie, Paris, France, ²⁰INSERM, U830, Paris, France, ²¹Université Paris-Descartes, Paris, France, ²²Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Centre Hospitalier Universitaire de Lyon/Centre Léon Bérard, Lyon, France, ²³Familial Cancer Centre, ²⁴Cancer Genetics Laboratory and ²⁵Kathleen Cunningham Foundation Consortium for Research into Familial Breast Cancer (kConFab), Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia, ²⁶Sir Peter MacCallum Department of Oncology and ²⁷Department of Pathology, University of Melbourne, Parkville, Victoria, Australia, ²⁸Department of Immunohematology and Transfusion Medicine and ²⁹Unit of Medical Oncology, Azienda Ospedaliera Papa Giovanni XXIII, Bergamo, Italy, ³⁰Dipartimento di Scienze Biomediche Sperimentali e Cliniche, Università di Firenze, Firenze, Italy, ³¹FiorGen Foundation for Pharmacogenomics, Sesto Fiorentino, Italy, ³²Institute of Medical Genetics, 'A. Gemelli' School of Medicine, Catholic University, Rome, Italy, ³³Dipartimento di Oncologia, Ematologia e Malattie dell'Apparato Respiratorio, Università di Modena e Reggio Emilia, Modena, Italy, ³⁴Immunology and Molecular Oncology Unit, Istituto Oncologico Veneto IOV – IRCCS, Padua, Italy, ³⁵Cancer Bio-Immunotherapy Unit and ³⁶Unit of Experimental Oncology 1, CRO Aviano National Cancer Institute, Aviano (PN), Italy, ³⁷Unit of Hereditary Cancers, IRCCS AOU San Martino – IST, Genoa, Italy, ³⁸Medical Genetics, University of Siena, Siena, Italy, ³⁹Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy, ⁴⁰Department of Molecular Medicine, Sapienza University of Rome, Rome, Italy, ⁴¹SC Medical Genetics, ASL8 Cagliari, Cagliari, Italy, ⁴²Unit of Pathology, Ospedale di Circolo, Varese, Italy, ⁴³Section of Genetic Oncology, University Hospital and University of Pisa, Pisa, Italy, ⁴⁴IRCCS Istituto Tumori 'Giovanni Paolo II', Molecular Genetics Laboratory, Bari, Italy, ⁴⁵Human Cancer Genetics Programme, Spanish National Cancer Centre (CNIO), Madrid, Spain, ⁴⁶Spanish Genotyping Centre (CEGEN), Madrid, Spain, ⁴⁷Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA, ⁴⁸Department of Human Genetics, ⁴⁹Department of Pathology and ⁵⁰Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands, ⁵¹Oncogenetics Group, Hospital Universitari de la Vall d'Hebron, Barcelona, Spain, ⁵²Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain, ⁵³Universitat Autònoma de Barcelona, Barcelona, Spain, ⁵⁴Department of Medical Oncology, Hospital Universitari de la Vall d'Hebron, Barcelona, Spain, ⁵⁵Center for Familial Breast and Ovarian Cancer, University Hospital of Cologne, Cologne, Germany, ⁵⁶Catalan Institute of Oncology – IDIBELL, Barcelona, Spain, ⁵⁷Clinical Genetics Service, ⁵⁸Department of Medicine and ⁵⁹Department of Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY, USA, ⁶⁰Division of Oncology, Department of Clinical Sciences, ⁶¹Department of Clinical Genetics, Laboratory Medicine, Office for Medical Services and ⁶²Department of Clinical Genetics, Lund University, Lund, Sweden

*To whom correspondence should be addressed at: IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, via Adamello 16, 20139 Milan, Italy. Tel: +39 2 57430 3812; Fax: +39 2 57430 3231; Email: paolo.peterlongo@ifom.eu

Abstract

Numerous genetic factors that influence breast cancer risk are known. However, approximately two-thirds of the overall familial risk remain unexplained. To determine whether some of the missing heritability is due to rare variants conferring high to moderate risk, we tested for an association between the c.5791C>T nonsense mutation (p.Arg1931*; rs144567652) in exon 22 of FANCM gene and breast cancer. An analysis of genotyping data from 8635 familial breast cancer cases and 6625 controls from different countries yielded an association between the c.5791C>T mutation and breast cancer risk [odds ratio (OR) = 3.93 (95% confidence interval (CI) = 1.28–12.11; $P = 0.017$)]. Moreover, we performed two meta-analyses of studies from countries with carriers in both cases and controls and of all available data. These analyses showed breast cancer associations with OR = 3.67 (95% CI = 1.04–12.87; $P = 0.043$) and OR = 3.33 (95% CI = 1.09–13.62; $P = 0.032$), respectively. Based on information theory-based prediction, we established that the mutation caused an out-of-frame deletion of exon 22, due to the creation of a binding site for the pre-mRNA processing protein hnRNP A1. Furthermore, genetic complementation analyses showed that the mutation influenced the DNA repair activity of the FANCM protein. In summary, we provide evidence for the first time showing that the common p.Arg1931* loss-of-function variant in FANCM is a risk factor for familial breast cancer.

Introduction

Breast cancer (OMIM #114480) is a common oncological disease that accounts for 23% of all malignancies in women and is estimated to cause 1 400 000 new cases and more than 450 000 deaths worldwide every year (1). It has been estimated that ~13% of all breast cancer cases have one or more affected relatives and that risks of breast cancer increase with greater numbers of affected relatives (2). This increased risk is also due to known germ-line susceptibility alleles including rare, high-risk loss-of-function variants predominantly found in BRCA1 and BRCA2 (3). In addition, 94 common single nucleotide polymorphisms (SNPs) have been identified that individually confer only a slightly increased risk of breast cancer, but combined in a multiplicative model account for ~16% of familial breast cancer risk (4).

BRCA1 and BRCA2 gene products contribute to cell homeostasis through the DNA damage response mediated by homologous recombination. Moreover, mutations in BRCA2 (also known as FANCD1) have been shown to cause Fanconi Anaemia (FA), a rare recessive disorder characterized by genomic instability, progressive bone marrow failure and predisposition to cancer. These genes encode proteins belonging to the FA pathway, which becomes activated in response to breaks in single- and double-stranded DNA. Monoallelic variants in several of these genes, including ATM, PALB2/FANCN and RAD51C/FANCO, have been detected in non-BRCA1 and BRCA2 familial breast cancer cases, but at a lower frequency in controls, consistent with moderate to high risks of breast cancer (5–7). The rare variants identified in these genes have a cumulative frequency in familial cases of 0.5–2%. However, with the exception of a few recurrent or founder mutations in specific populations, each of these mutations is generally very rare with many reported in single families. In contrast, few rare truncating and pathogenic missense variants have been found in CHEK2 (8), with much of the risk attributed to this gene explained by the single moderate-penetrance founder allele, c.1100delC (9).

Recent studies have underlined the challenges in identifying new breast cancer predisposition genes. Exome sequencing in families followed by gene re-sequencing in additional cases and controls have provided conflicting results for XRCC2 (10,11), and inconclusive results for FANCC and BLM (12), raising questions about the statistical power of these studies (13). Similarly, the evidence that SLX4, an FA gene, is associated with breast cancer risk is limited, given that the analysis of large numbers of familial cases identified only three inactivating variants (14–16).

Screening for risk-associated mutations in BRCA1 and BRCA2 is commonly used in clinical practice to identify at-risk individuals and to direct them towards specific surveillance programmes or risk reduction options. By including additional breast cancer predisposition genes in gene panels analysed by next-generation sequencing, risk prediction can be performed in a larger fraction of individuals at a reduced cost with rapid turnaround time. With the goal of identifying new risk-associated genes, we and others previously performed exome sequencing in multiple-case breast cancer families (17). One of the findings of that study was a single proband heterozygous for the c.5791C>T variant (rs144567652) in FANCM, another gene involved in the FA pathway. The variant was predicted to introduce a stop codon (TGA) in exon 22, causing the loss of 118 amino acids from the C-terminus (p.Arg1931*). A subsequent case-control study detected the mutation in 10 of 3409 (0.29%) familial cases without known mutations in BRCA1 and BRCA2 and in 5 of 3896 (0.13%) controls from different national studies. The estimated odds ratio (OR) was 2.29 [95% confidence interval (CI) = 0.71–8.54; $P = 0.13$]. In an effort to establish the significance of this estimate (17), a further analysis in a larger cohort was performed.

Results

Association with breast cancer risk

We investigated the c.5791C>T mutation in a large series of familial cases without known mutations in BRCA1 and BRCA2 and in a comparable set of control individuals from Italy, France, Spain, Germany, Australia, USA, Sweden and The Netherlands. The mutation was found in 18 of 8635 (0.21%) cases (pedigrees are shown in Supplementary Material, Fig. S1) and in 4 of 6625 (0.06%) controls (Table 1) giving a statistically significant association with breast cancer risk with an age-adjusted OR of 3.93 (95% CI = 1.28–12.11; $P = 0.017$). The c.5791C>T mutation is rare and we observed a large variation in allele frequency in cases and controls across studies. To control for population stratification, we performed a meta-analysis, including only studies in which mutation carriers were detected in both cases and controls (Italy, France and Australia). Starting from the ORs and their 95% CIs obtained from a univariate logistic model within each country, we obtained a pooled OR = 3.67 (95% CI = 1.04–12.87; $P = 0.043$) (Table 2). A second meta-analysis was performed by exploiting all the available data. We implemented an exact conditional logistic regression model including 'country' as a random covariate

Table 1. Number and frequency of mutation carriers and non-carriers in cases and controls

Geographical group	Country	Cases			Controls		
		Carriers	Non-carriers	Freq%	Carriers	Non-carriers	Freq%
South/Western Europe	Italy	6	2209	0.27	1	1483	0.07
	France	5	1570	0.32	1	1323	0.08
	Spain	3	751	0.40	0	286	NA
	All	14	4530	0.31	2	3092	0.06
Non-South/Western Europe	Germany	0	1636	NA	1	1899	0.05
	Australia	3	1235	0.24	1	1164	0.09
	USA	0	517	NA	0	322	NA
	Sweden	0	484	NA	0	0	NA
	The Netherlands	1	215	0.46	0	144	NA
	All	4	4087	0.10	2	3529	0.06
All populations	Total	18	8617	0.21	4	6621	0.06

NA, not applicable.

Table 2. Meta-analysis of the study results from countries with mutation carriers in both cases and controls

Country	Cases			Controls			OR	95% CI	P-value
	Carriers	Non-carriers	Freq%	Carriers	Non-carriers	Freq%			
Italy	6	2209	0.27	1	1483	0.07	4.03	0.48–33.47	0.197
France	5	1570	0.32	1	1323	0.08	4.21	0.49–36.10	0.189
Australia	3	1235	0.24	1	1164	0.09	2.82	0.24–27.13	0.369
Pooled	14	5014	0.28	3	3970	0.08	3.67	1.04–12.87	0.043

OR, odds ratio; CI, confidence interval.

in order to control for population stratification and for the absence of variant carriers in some countries (Sweden and USA) [OR = 3.330 (95% CI = 1.087–13.615; P = 0.0320)].

Expression of the mutant allele

To verify the functional consequences of the c.5791C>T mutation, we first performed reverse transcriptase-polymerase chain reaction (RT-PCR) in lymphoblastoid cell lines (LCLs) derived from two mutation carriers. Amplifying a product spanning exons 22 and 23, sequence analyses revealed very low levels of the mutated transcript compared with corresponding normal mRNA. Treatment with a protein synthesis inhibitor, cycloheximide, did not alter mutant transcript levels (Fig. 1), suggesting that the effect was probably not related to nonsense-mediated decay (NMD), but rather to a defect in mRNA splicing itself.

Effect on the mRNA splicing

The occurrence of exonic mutations affecting pre-mRNA splicing is well documented in many human disease genes (18). These include nonsense mutations, a phenomenon referred to by some authors as ‘nonsense-associated altered splicing’ (19). Therefore, to assess the impact of c.5791C>T on splicing regulatory binding sites controlling exon definition, we performed information theory-based mutation analysis, using the Automated Splice Site and Exon Definition Analysis (ASSEDA) server (20). The variant was predicted to create a strong binding site [information content (R_i) = 4.6 bits] for the splicing factor hnRNP A1 at position c.5790_5795 (Fig. 2A) (21). Exon definition analysis suggested that the creation of this site would completely suppress exon recognition ($R_{i, total}$ from 3.5 to –2.5 bits), predicted to result in

exon skipping. This was confirmed by RT-PCR, using forward and reverse primers in exons 21 and 23 which detected two amplification products (Fig. 2B). The upper band derived from the full-length transcript (Supplementary Material, Fig. S2), whereas the lower band derived from an aberrant transcript lacking the entire exon 22 (c.5717_6008del292) (Fig. 2C). This exon skipping is predicted to encode a protein that incorporates 11 additional residues and to lead to a premature termination of translation that results in the loss of 132 amino acids from the FANCM C-terminus (p.Gly1906Alafs12*).

Skipping of exon 22 is mediated by hnRNP A1

A pull-down experiment with HeLa cell extracts followed by western blot analysis showed that the hnRNP A1 protein specifically bound to RNA oligonucleotides spanning FANCM position r.5779_5804 and carrying the r.5791C>U mutation, whereas a very weak interaction was observed with the corresponding normal oligonucleotide (Fig. 2D). These results are in agreement with the outcomes of the *in silico* analyses and provide evidence that the mechanism through which the c.5791C>T mutation causes exon 22 skipping is mediated by the binding of hnRNP A1 protein.

DNA repair activity-based functional studies

We then checked whether the c.5791C>T mutation ($\Delta 22$) affects FANCM activity in DNA repair, by genetic complementation of the following FA-associated cell phenotypes: hypersensitivity to mitomycin C (MMC)- and diepoxybutane (DEB)-induced chromosome fragility. As an internal control, we used a previously described nonsense mutation (p.S724X; rs137852864) (22) that leads to a premature stop codon (stop). Mutant cDNAs were

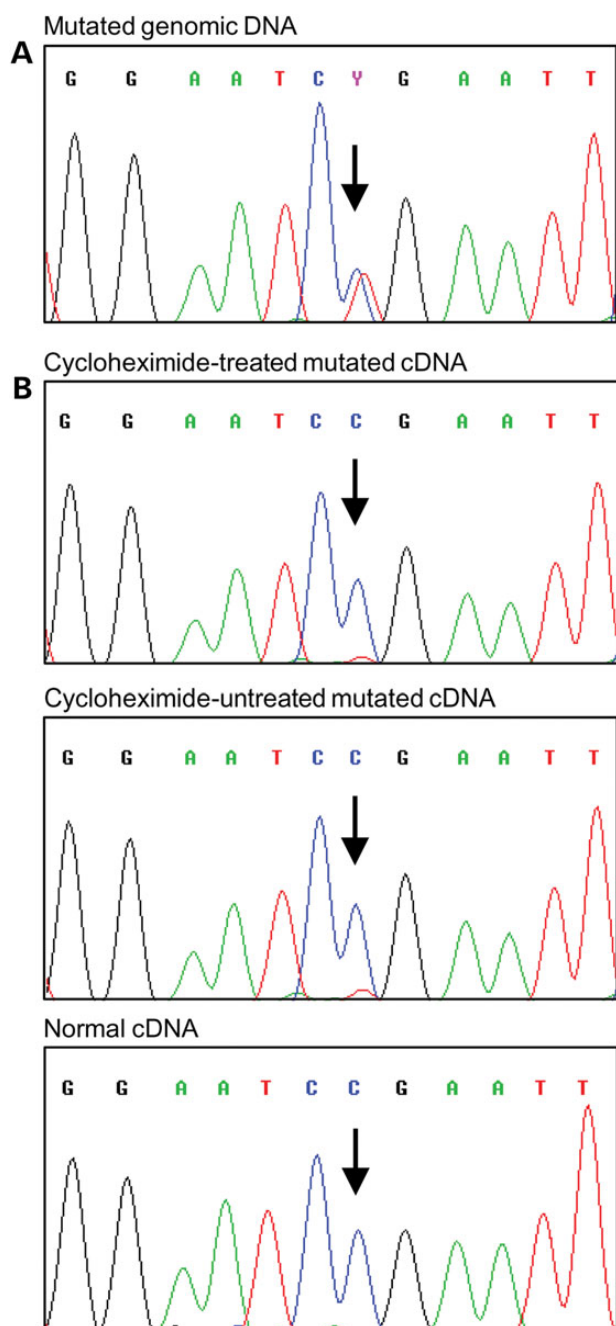


Figure 1. Sequencing analysis of the FANCM gene and transcript. (A) Genomic DNA fragment PCR amplified using both primers in exon 22 from an LCL carrying the c.5791C>T mutation. (B) cDNA fragment amplified by PCR using a forward primer in exon 22 and a reverse primer in exon 23. A strong reduction in the expression of the mutant allele was observed in both cycloheximide treated and untreated cells. cDNA from a non-carrier individual was used as a control. The position of the mutation is indicated by the arrows. Identical results were observed in an additional mutated LCL.

generated by site-directed mutagenesis and cloned into lentiviral vectors to stably transduce *Fancm*^{-/-} immortalized mouse embryonic fibroblast (MEFs). Wild-type, $\Delta 22$ and stop alleles were expressed at similar levels in infected cells (Supplementary Material, Fig. S3). As expected, wt FANCM but not the prematurely truncated forms rescued MMC hypersensitivity of *Fancm*^{-/-} MEFs (Fig. 3A). Similar results were observed in a chromosome fragility

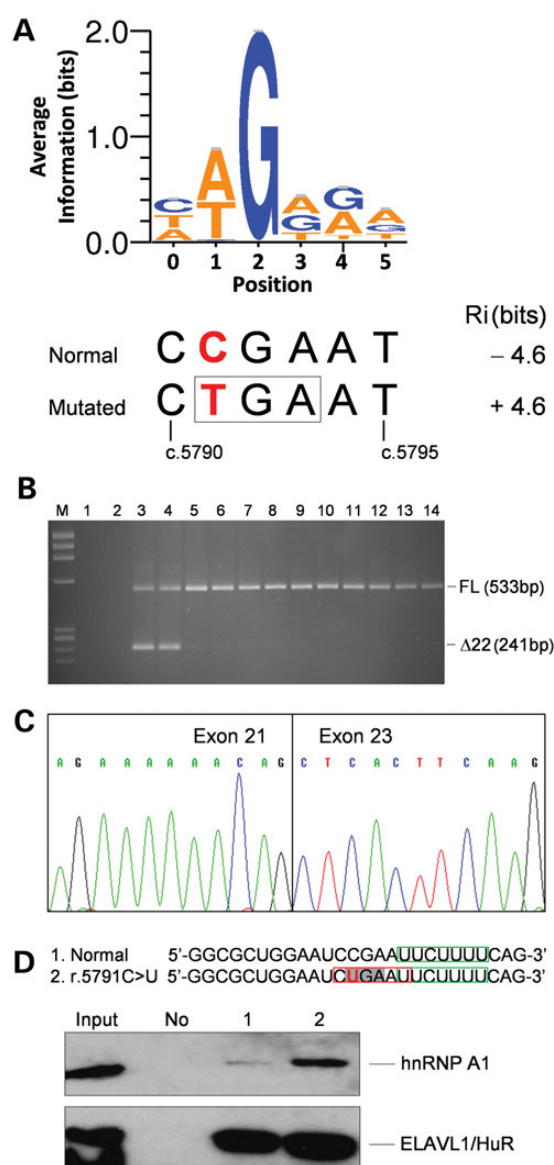


Figure 2. Analysis of the effect of the FANCM c.5791C>T mutation on RNA. (A) Sequence logo of hnRNP A1 binding sites generated as described in Materials and Methods. The opal codon (TGA, boxed) introduced by the FANCM c.5791C>T mutation (in bold red) is contained at positions 1–3 of the hnRNP A1 binding-site encompassing nucleotides c.5790_5795. The hnRNP A1 binding-site strength computed by the ASSEDA software for the normal and mutated sequences is reported. (B) Agarose gel electrophoresis of the RT-PCR products using a forward primer in exon 21 and a reverse primer in exon 23. M, molecular marker (Φ X-174 HaeIII digested); 1, no template as a negative control for PCR; 2, genomic DNA as a negative control for the specificity of cDNA amplification; 3 and 4, cDNAs from LCLs carrying the c.5791C>T mutation; 5–14, cDNAs from LCLs derived from 10 mutation negative individuals, used as reference controls. The sizes of the full-length (FL) and Δ exon22 ($\Delta 22$) transcripts are indicated. (C) Sequence of the aberrant band excised from the gel showing the skipping of the entire exon 22. (D) Western blot analysis of biotin RNA–hnRNP A1 protein pull down using a goat polyclonal antibody. The sequence of the used RNA oligonucleotides encompassing FANCM positions r.5779_5804 is reported, with the r.5791C>U mutation in bold red, the opal codon enlightened in light grey and the predicted hnRNP A1 binding site created by the mutation boxed in red. As a control for the pull-down efficiency and specificity, we used an antibody against the ELAVL1/HuR protein for which a binding site, boxed in green, is predicted in both RNA oligonucleotides. Input, 10% of total HeLa cell line extract used in the pull-down assay. No, no RNA used as a negative control; 1, normal RNA; 2, RNA carrying the r.5791C>U mutation. The results shown here are representative of two independent experiments.

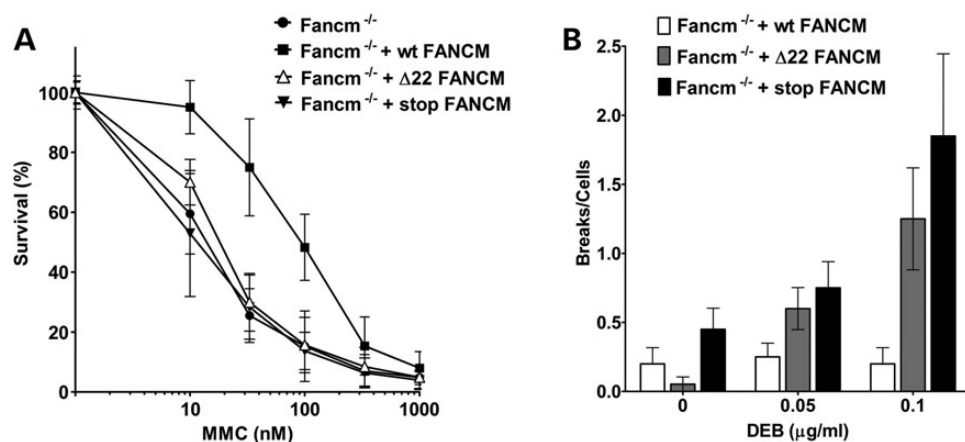


Figure 3. Functional studies of the FANCM mutation. (A) Analysis of cellular MMC sensitivity. MEFs expressing Δ22-FANCM allele (Fancm^{-/-} + Δ22 FANCM) are more sensitive to MMC than the cell expressing the wt FANCM allele (Fancm^{-/-} + wt FANCM). Not transduced MEFs (MEF Fancm^{-/-}) and MEF Fancm^{-/-} expressing a FANCM with loss-of-function mutation p.S724X (Fancm^{-/-} + stop FANCM) are used as controls (N=3; error = standard deviation). (B) Chromosome fragility induced by DEB treatment. Fancm^{-/-} + Δ22 FANCM and Fancm^{-/-} + stop FANCM cells show higher chromosome fragility than Fancm^{-/-} + wt FANCM. Twenty metaphases were analysed for chromosome breaks. Results are represented as mean number of breaks per cells and the error bars are SEM.

assay (Fig. 3B). These observations lead to the conclusion that the Δ22 form of FANCM is deficient in DNA repair of MMC- and DEB-induced stalled replication forks.

Discussion

Genotyping of the c.5791C>T variant from FANCM in 8635 familial cases with no mutations in BRCA1 and BRCA2 and in 6625 control individuals from Italy, France, Spain, Germany, Australia, USA, Sweden and The Netherlands indicate that the variant is associated with risk of breast cancer with OR of 3.93. The association between the variant and breast cancer risk is reinforced if the data from the present case-control analysis are combined with those of Gracia-Aznarez et al. (17). Overall, the variant was detected in 28 of 12 044 (0.23%) familial cases and 9 of 10 521 (0.09%) controls, corresponding to a naive OR, adjusted for the 'study' covariate, of 2.83 (95% CI = 1.33–6.01; P = 0.007). It has to be noted that these studies were based on cases with positive family history for breast cancer and/or disease early onset. These selected cases are likely to be enriched in predisposing genetic factors. Consequently, the ORs observed here could be higher than those expected in unselected cases and population controls.

Population stratification may occur when a rare mutation is tested in individuals from different countries. We took into consideration this critical issue by performing two meta-analyses. The first was based only on studies with carriers in both cases and controls, whereas the second exploited all the individual data. Both analyses supported the c.5791C>T mutation as a breast cancer risk factor (OR = 3.67 and 3.33, respectively), although with borderline statistical significance. However, these analyses do not completely guard against stratification effect. Hence, these results need to be taken with caution. In this light, genotyping of additional variants in much larger series of unselected breast cancer cases and matched controls will be needed to confirm FANCM as a breast cancer susceptibility gene.

The functional characterization of the mRNA transcript derived from the c.5791C>T allele shows that the variant causes the skipping of exon 22 introducing a premature stop codon. In addition, genetic complementation assays revealed the mutated protein lacks DNA repair activity. These results support that the

FANCM c.5791C>T mutation is pathogenic. Of note, our observations at the mRNA level emphasize the notion that, although nonsense mutations are usually considered as inherently deleterious, transcript analyses are required for a precise assessment of actual functional consequences.

Interestingly, while this article was in preparation, another FANCM C-terminus truncating mutation, the c.5101C>T (p. Q1701*), was detected by the exome sequencing of 11 Finnish breast cancer families (23). The c.5101C>T mutation was shown to be significantly more frequent in Finnish breast cancer cases compared with controls (OR = 1.86; 95% CI = 1.26–2.75; P = 0.0018), to have higher effect in cases with family history (OR = 2.11; 95% CI = 1.34–3.32; P = 0.0012)—although incomplete co-segregation with the disease was observed among most of the families—and with the stronger effect (OR = 3.56; 95% CI = 1.81–6.98; P = 0.0002) in mutation carriers affected with triple-negative (oesrogen receptor-, progesterone receptor- and HER2-negative) breast cancer (TNBC). In addition, the age at breast cancer diagnosis was not different between variant carriers and non-carriers and the variant allele was not subjected to mRNA NMD (23). Among the cases we studied, there were no significant differences in age at breast cancer diagnosis in the 18 variant carriers versus the non-carriers (data not shown). Moreover, because variant carriers with TNBC were not found, it was not possible to associate the c.5791C>T mutation with this specific tumour subtype (see Supplementary Material, Table S1). Our data, together with the observations by Kiiski et al., support the notion that loss-of-function mutations of FANCM are moderate breast cancer risk factors. Additional studies, such as modified segregation-analysis in families, are warranted to provide age-specific risk estimates for FANCM mutation carriers.

The 1100delC allele of CHEK2 has a number of properties similar to c.5791C>T in FANCM. This allele has been shown to be associated with moderate risk and has been found in many countries with a frequency that is higher in North-Eastern Europe and decreases in Southern Europe (9). Taking into consideration the geographic origin of the individuals included in our study, we observed that the frequency of the mutation carriers in South-Western European countries (16/7638 = 0.21%) was higher, although with borderline statistical significance (Fisher's exact test, P = 0.052), than that observed in all other countries

evaluated (6/7622 = 0.08%) (Table 1). This suggests that this FANCM mutation has a frequency gradient that is opposite of that reported for the CHEK2 1100delC. Nevertheless, the relatively small numbers of carriers observed suggest that further analyses in specific populations would be worthwhile. For example, data from the Exome Aggregation Consortium (ExAC), Cambridge, MA (<http://exac.broadinstitute.org>) (March 2015, accessed), that is collecting exome sequencing data from individuals included in various disease-specific and population genetic studies, indicate that c.5791C>T occurs in the Finnish population with a carrier frequency of nearly 1%.

The FA pathway is generally subdivided into upstream proteins, assembling into the 'core complex' and 'downstream effectors'. The 'core complex' ubiquitinates the FANCI-FANCD2 complex and this activates the pathway coordinating the action of the downstream effectors. The latter include FANCD1/BRCA2, FANCN/PALB2, FANCF/BRIP1 and FANCO/RAD51C that are required for DNA repair by homologous recombination and are all breast cancer genes (24). This has suggested that the involvement of FA genes in breast cancer susceptibility could be limited to the FA downstream effectors (24,25). The FANCM protein has different functional domains, including a DEAH translocase domain at the N-terminus, and domains for interaction with proteins mediating DNA binding between amino acids 675 and 790 for interaction with MHF1 and MHF2, and at the C-terminus beyond amino acid 1799, for interaction with FAAP24 (26). Moreover, FANCM-FAAP24-MHF1-MHF2 acts as an independent 'anchor complex' that recognizes the damage caused by interstrand cross-linking agents and recruits the FA core complex (26). Furthermore, FANCM is not essential for complete ubiquitination of the FANCI-FANCD2 complex (24,26). Finally, the FANCM/MHF complex has a translocase activity that is independent of the core complex proteins (27). Hence, FANCM has a direct activity in maintaining the DNA integrity that can be independent of the FA pathway. In this light, it is possible that the increased risk for breast cancer conferred by the c.5791C>T variant is due to a direct impairment of the DNA damage response as for the FA downstream effectors.

BRCA1 and BRCA2 are established breast cancer risk factors with high penetrance and PALB2 was also recently shown to confer a high breast cancer risk (28). Although these three genes encode for downstream effectors of the FA pathway, their proteins are also involved in other DNA damage responses, including double-strand break repair. Thus, BRCA1, BRCA2 and PALB2 exert DNA damage response functions that one can speculate to be of greater magnitude than FANCM. Consequently, susceptibility to breast cancer is expected to be higher in BRCA1, BRCA2 or PALB2 mutation carriers and lower in carriers of FANCM mutations, which is in agreement with our data on the c.5791C>T mutation and similar data on the c.5101C>T mutation (23). By analogy, the abrogation of BRCA1, BRCA2 and PALB2 functions on the one hand, and the abrogation of FANCM function on the other, seem to impact FA differentially. It is known that biallelic BRCA2 and PALB2 mutations cause FA, and recently, initial evidence of the involvement of BRCA1 mutations in the disease have been documented. BRCA1 biallelic mutations were found in a woman showing anomalies consistent with a FA-like disorder, which supports BRCA1/FANCS as a novel FA gene (29). In contrast, the role of FANCM in FA is questionable. The only FA patient reported so far with truncating FANCM mutations also carried deleterious mutations in FANCA (22,30). Moreover, individuals homozygous for the truncating FANCM mutations c.5101C>T and c.5791C>T did not present with FA

(31) indicating that, at present, FANCM cannot be considered to cause FA.

In conclusion, based on the analysis of large sets of cases and controls and functional observations, our study provides evidence that the FANCM c.5791C>T is a novel risk factor for familial breast cancer.

Materials and Methods

Study populations

Twenty-five case-control cohorts were included in the study through a collaboration call circulated among the COMPLEXO participants (32). Centre or study details, number and description of cases and controls are reported in Supplementary Material, Table S2. The cases included in this study were (i) affected with breast cancer at age ≥ 18 , (ii) eligible to BRCA1- and BRCA2-mutation testing based on breast and/or ovarian cancer family history (at least one first- or second-degree female relatives with either tumours), or because affected with early onset (≤ 40 years) or bilateral (≤ 50 years) breast cancer and (iii) negative to BRCA1- and BRCA2-mutation test. A few of the centres or studies contributing to this article used slightly different inclusion criteria, and these are described in Supplementary Material, Table S2. In all cohorts, cases and controls were female Caucasians recruited in the same area. All individuals included in the study signed an informed consent to the use of their biological samples for research purposes. The participation to this study was approved by ethical committees or review boards of the participant centres or studies.

Mutation genotyping

Details of genotype analyses for FANCM (NG_007417.1, NM_020937.2) c.5791C>T mutations are reported in Supplementary Material, Table S2. For most studies, cases and controls were genotyped at coordinating centre (IFOM, Milano) by custom TaqMan SNP genotyping assay (Life Technologies) using the following primers and probes. Forward primer: 5'-AGCCTGCTGACTACCTTAATTGG-3'; reverse primer: 5'-CTTTAGCAAATCTGGGTTTCTTCT-3'; probe 1: 5'-TGAAAAGAATTCGGATTCC-3'; probe 2: 5'-TGAAAAGAATTCAGATTCC-3'. In every 96-well plate, at least one positive control and two blank controls were included. The remaining samples were genotyped at local centres using TaqMan assay or high-resolution melting. All positive samples were confirmed by double-strand Sanger sequencing. Two studies (SWE-BRCA and MAYO) provided genotyping data from previous next-generation sequencing studies.

Statistical analyses

Logistic regression analysis was used to test the association between mutation frequencies and risk of breast cancer (33). Age was included in the model as adjustment covariate and the adjusted ORs and its 95% CIs were estimated. A meta-analysis considering only countries in which mutation carriers were observed in both cases and controls was performed based on mixed models (34) starting from single-study estimates. A meta-analysis exploiting all individual data was performed using an exact conditional logistic regression model (35). The statistical analyses were performed with the SAS software (Version 9.2; SAS Institute Inc., Cary, NC).

Information theory-based mutation analyses

The ASSEDA server (<http://mutationforecaster.com>) has been developed to predict the molecular phenotype of putative

splicing mutations (20) and implemented to take into account their effect on splicing factor binding sites. In particular, CLIP-seq libraries for hnRNP A1 (36) were used to derive information theory-based position weight matrix (PWM), depicted in Figure 2A. PoWeMaGen software, which uses Bipad (37) to generate minimum entropy alignments, generates a series of potential binding-site models over a range of input parameters. To mitigate against phasing the alignment on natural splice sites instead of adjacent hnRNP A1 binding sites, models were built from shorter sequences, ranging in lengths between 18 and 25 nucleotides (nt). The optimal model was determined by maximizing incremental information by varying binding-site length (6–10 nt), number of Monte Carlo cycles (250–5000) and allowing either zero or only one site per sequence (OOPS). The model with the highest average information used a maximum fragment length of 18 nt, 1000 Monte Carlo cycles, OOPS and a single-block binding-site length of 6 nt. This sequence is frequently present in sites cross-linked to hnRNP A1 protein (34). Of the 140 431 hnRNP A1 binding sites used to create the information theory-based model, the wild-type sequence, CCGAAT, is not represented, and the mutant site, CTGAAT, occurred 716 times. The model was validated with known hnRNP A1 binding sites and splicing affecting mutations (38–41). The effects of mutations at hnRNP A1 sites on exon definition were determined from the total information content ($R_{i,\text{total}}$), by incorporating changes in the strengths of these sites, corrected for the gap surprisal, which represents the distance between the hnRNP A1 site and the natural splice site. Gap surprisal values were determined by scanning the genome for hnRNP A1 sites with the PWM, and then determining the frequency of each interval length between known natural sites and the nearest hnRNP A1 site, separately for exons and introns. Differences between the natural and mutated exon $R_{i,\text{total}}$ values correspond to changes in the abundance of the respective isoforms and can predict exon skipping. The calculation is carried out by the ASSEDA server (20). Exon definition analysis was validated for a set of mutations that affect hnRNP A1 binding-site strength (39,40).

Cell lines

Epstein–Barr virus-immortalized human LCLs were established from peripheral blood derived from 2 carriers of the c.5791C>T mutation and 10 normal controls (42). LCLs were maintained in RPMI 1640 medium supplemented with 15% foetal bovine serum plus 1% penicillin–streptomycin. The HeLa cell line was cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum plus 1% penicillin–streptomycin at 37°C in a humidified 5% CO₂ atmosphere. The HeLa cell line was authenticated by short tandem repeat analysis using the kit GenePrint10 kit (Promega). *Fancm*^{-/-} immortalized MEFs (43) were a kind gift of Dr H. Te Riele from The Netherlands Cancer Institute, Amsterdam. MEFs were cultured in DMEM 10% fetal calf serum supplemented with antibiotics. All the cell lines used in this study were routinely checked for mycoplasma contamination using the PCR Mycoplasma Detection Set (Takara) or the MycoAlert™ Mycoplasma Detection Kit (Lonza).

Transcript analyses

Potential degradation of unstable transcripts containing premature termination codons via nonsense-mediated mRNA decay (NMD) was prevented by growing LCLs in the presence of cycloheximide (100 µg/ml) for 4 h. Total RNA was purified from LCLs using the Nucleospin RNA II (Macherey-Nagel) and the cDNA

was synthesized using random primer and Maxima H Minus Enzyme (Thermo Scientific), according to the manufacturers' protocols. cDNAs were PCR amplified using the following primers. Exon 21, forward: 5'-CAAGTTCATTGAGCAGATCCAG-3'; exon 22, forward: 5'-ACATCAAGGATGTTTAGGA-3'; exon 22, reverse: 5'-GTGCCTCACTTTTATTACTA-3'; exon 23, reverse: 5'-CCCATCTTGAGCAGCTTGA-3'. Amplification products were visualized on agarose gel stained with ethidium bromide and directly characterized by Sanger sequencing. Alternatively, single PCR fragments were excised from the gel, purified using the Wizard® SW Gel and PCR Clean-Up System (Promega) and sequenced.

Biotin RNA–protein pull-down assay

Protein extraction was performed starting from $\sim 5 \times 10^6$ HeLa cells. These cells were harvested by centrifugation at 4°C for 5 min, washed twice with 1× phosphate buffered saline (PBS) and lysed in lysis buffer [25 mM Tris (pH 7.4), 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid, 1% Nonidet P-40, 5% glycerol] containing protease inhibitor (Sigma–Aldrich), on ice for 30 min. Following centrifugation, the protein concentration was determined by Bradford assay (Bio-Rad). Mutated and normal RNA oligonucleotides were biotinylated at the 3' end using the RNA 3' End Desthiobiotinylation Kit (Thermo Scientific Pierce), according to the manufacturer's instructions. For each binding reaction, 50 pmol of biotinylated RNA oligonucleotides were coupled to 50 µl of Streptavidin Magnetic Beads (Thermo Scientific Pierce) and incubated with an equal amount of HeLa cell lysates in 1× protein–RNA binding buffer [0.2 M Tris (pH 7.5), 0.5 M NaCl, 20 mM MgCl₂, 1% Tween-20 detergent], for 2 h at 4°C with agitation. The bound proteins were eluted from the magnetic beads by incubating with 50 µl of biotin elution buffer (Thermo Scientific Pierce) for 30 min at 37°C with agitation. The eluted proteins were subjected to 4–15% sodium dodecyl sulphate–polyacrylamide gel electrophoresis gradient gel and visualized by western blotting using a goat polyclonal antibody against hnRNP-A1 (#sc-10029, Santa Cruz Biotechnology) or a mouse monoclonal antibody against ELAVL1/HuR (#1862775, Thermo Scientific Pierce). The binding site for the ELAVL1/HuR protein was identified with ASSEDA (<http://splice.uwo.ca/logos.html>).

Plasmids used for functional studies

The doxycycline-inducible lentiviral vector pLVX-TRE3G-FANCM was kindly provided by Dr N. Ameziane (Vrije Universiteit Medical Center, Amsterdam, The Netherlands) and mutated by site-directed mutagenesis with the QuickChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies), as previously described (42) and using the following PAGE purified mutagenic primers. $\Delta 22$ primer 1: 5'-GAAAAGGACAGAGAAAAAACAGCTCACTTCAAGAAATCTCCATG-3', $\Delta 22$ primer 2: 5'-CATGGAGATTCTTTGAAGTGAGCTGTTTTTCTCTGTCTTTTC-3', c.2171C>A primer 1: 5'-TGAGGAAAACAACCAGCTCAAGAATAAACCACTGGAATTC-3' and c.2171C>A primer 2: 5'-GAATCCAGTGGTTATTCTTGAGCTGGTTTGTCTTCTCA-3'. The sequences of all FANCM constructs were verified by Sanger sequencing to confirm that they only bore the intended mutations.

Lentiviral particles production and cell transduction

To prepare lentiviral particles, 5×10^6 HEK-293 T cells were plated into 10 cm dishes. The next day, medium was changed with a fresh one containing 30 µM Chloroquine (Sigma) and cells were transfected with the lentiviral expression vectors and the helper

plasmids (PAX and ENV) using the CalPhos Mammalian Transfection Kit (Clontech). The medium was changed 24 h after transfection, and 24 h later, the lentivirus-containing supernatant was collected and filtered through a 0.45 µm filter (Millipore). Additional supernatant was collected after additional 24 h, filtered and pooled with the initial one. Pooled supernatants were centrifuged in a Beckman JS-24.38 rotor at 19 500 rpm for 1.5 h at 4°C. Pellets were resuspended in PBS (50 µl of PBS/10 ml of supernatant) and stored at -80°C. Sixty thousands of *Fancm*^{-/-} MEFs per well were seeded in a 12-well plate. After 24 h, cells were infected with 20 µl of concentrated viral supernatant in the presence of 1 µg/ml polybrene (Millipore). Twenty-four hours later, infected cells were selected with puromycin (2.5 µg/ml). Transgene expressions were checked by real-time PCR.

MMC sensitivity test

Twenty-five thousand cells of each cell line were seeded in 2 ml of complete medium supplemented with 2 µg/ml doxycycline in 12 wells of 6-well plates. The next day, MMC was added at the indicated doses and the cell sensitivity was evaluated 72 h after cell cultures were washed with PBS. Cells were collected in a volume of 300 µl of trypsin and 700 µl of complete medium and counted with a Z2™ coulter counter (Beckman Coulter) (44).

Chromosome fragility test

Two hundred thousand cells were seeded in complete medium supplemented with 2 µg/ml doxycycline. Twenty-four hours later, DEB was added at the indicated concentrations, and metaphase spreads were then harvested 3 days later as it follows. Colcemid™ (Sigma) was added at 0.1 µg/ml final concentration and after 2 h, cells were trypsinized, washed in PBS, the pellet was resuspended in hypotonic solution (0.075 M KCl) and incubated for 25 min at 37°C. Cells were then washed three times with methanol:acetic acid (4:1) and the cell suspension was dropped on microscope slides and Giemsa stained. Twenty metaphase cells from each DEB concentration were scored for chromosome breakage after image capture using the Metafer Slide Scanning Platform from Metasystems (45).

Supplementary Material

Supplementary Material is available at HMG online.

Acknowledgements

The authors wish to thank all the individuals who took part in this study. We also wish to thank Antonis Antoniou (University of Cambridge) for helpful discussion, Claudia Foglia (INT) for DNA preparation, Fernando Ravagnani (INT) for recruitment of blood donors and Donata Penso (INT) for cell lines authentication; Roser Pujol and Maria Jose Ramirez (CIBERER/UAB) for cytogenetic analysis; Victoria Fernandez (CNIO) for the genotyping study; Gabriel Capella (Hereditary Cancer Program ICO) for the clinical and molecular work done in identifying and genotyping breast cancer patients; Peter Bugert (University of Heidelberg) for collecting control samples; Rongxi Yang and Katharina Mattes for supporting the genotyping; all the collaborating cancer clinics of the French National Study GENESIS (GENE SISTers), M. Marcou, D. Le Gal, L. Toulemonde, J. Beauvallet, N. Mebirouk, E. Cavaciuti, A. Fescia (genetic epidemiology platform The PIGE, Plateforme d'Investigation en Génétique et Epidemiologie), C. Verny-Pierre and L. Barjhoux (Biological Resource Centre); H. te Riele and

N. Ameziane for sharing materials; all the collaborating centers of the Swedish national study SWEA (SWE-BRCA Extended Analysis).

Conflict of Interest statement. P.K.R. is the inventor of US Patent 5,867,402 and other patents pending, which predict and validate mutations. He is one of the founders of Cytognomix, Inc. (London, Canada), which creates software based on this technology. All the others authors declare no conflict of interest.

Funding

This study was partially supported by funds from 'Ricerca Finalizzata - Bando 2010' from Ministero della Salute, Italy, to P.P. and C.T., Associazione Italiana per la Ricerca sul Cancro (AIRC-IG 12821) to P.P., (IG 12780) to L.O., (IG 5706) to L.V., and ("5xmille" n. 12237) to A.F., Italian citizens who allocated a 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori to S.M. and P.R., of the CRO Aviano National Cancer Institute to A.V. and of the Istituto Oncologico Veneto IOV - IRCCS to M.M., according to Italian law and Fondazione Italiana per la Ricerca sul Cancro' (FIRC)—triennial fellowship 'Armanda e Enrivo Mirto' to I.C. and triennial fellowship 'Mario e Valeria Rindi' to V.S.; National Institutes of Health (NIH, CA128978, CA116167 and CA176785) and the Breast Cancer Research Foundation to F.J.C.; Spanish Network on Rare Diseases (CIBERER) and the PI12/00070; the Spanish Carlos III Health Institute (FIS project PI12/02585) to O.D.; Asociación Española Contra el Cáncer, Spanish Health Research Foundation, Ramón Areces Foundation, Catalan Health Institute and Autonomous Government of Catalonia (ISCHIRETIC RD12/0036/008, PI10/01422, PI12/01528, PI13/00285, 2009SGR290 and 2009SGR283); the Swedish Cancer Society, Berta Kamprad Foundation, Gunnar Nilsson Foundation, BioCARE and the Swedish Society of Medicine; UNICANCER, the Ligue Nationale Contre le Cancer and the French National Institute of Cancer (INCa); MINECO (SAF2012-31881), ICREA Academia, Generalitat de Catalunya (SGR0489-2009, SGR317-2014) and European Regional Development FEDER Funds to J.S.; Canadian Breast Cancer Foundation, Canadian Foundation for Innovation, Canada Research Chairs Secretariat and the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant 371758-2009) to P.K.R.

References

1. Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C. and Parkin, D.M. (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer*, **127**, 2893–2917.
2. Collaborative Group on Hormonal Factors in Breast Cancer. (2001) Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet*, **358**, 1389–1399.
3. Mavaddat, N., Antoniou, A.C., Easton, D.F. and Garcia-Closas, M. (2010) Genetic susceptibility to breast cancer. *Mol. Oncol.*, **4**, 174–191.
4. Michailidou, K., Beesley, J., Lindstrom, S., Canisius, S., Dennis, J., Lush, M.J., Maranian, M.J., Bolla, M.K., Wang, Q., Shah, M. et al. (2015) Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat. Genet.*, **47**, 373–380.
5. Renwick, A., Thompson, D., Seal, S., Kelly, P., Chagtai, T., Ahmed, M., North, B., Jayatilake, H., Barfoot, R., Spanova, K. et al. (2006) ATM mutations that cause ataxia-telangiectasia

- are breast cancer susceptibility alleles. *Nat. Genet.*, **38**, 873–875.
6. Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A., Reid, S., Spanova, K., Barfoot, R., Chagtai, T. et al. (2007) PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.*, **39**, 165–167.
 7. Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D., Freund, M., Lichtner, P., Hartmann, L., Schaal, H. et al. (2010) Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat. Genet.*, **42**, 410–414.
 8. Le Calvez-Kelm, F., Lesueur, F., Damiola, F., Vallee, M., Voegelé, C., Babikyan, D., Durand, G., Forey, N., McKay-Chopin, S., Robinot, N. et al. (2011) Rare, evolutionarily unlikely missense substitutions in CHEK2 contribute to breast cancer susceptibility: results from a breast cancer family registry case-control mutation-screening study. *Breast Cancer Res.*, **13**, R6.
 9. Weischer, M., Bojesen, S.E., Ellervik, C., Tybjaerg-Hansen, A. and Nordestgaard, B.G. (2008) CHEK2*1100delC genotyping for clinical assessment of breast cancer risk: meta-analyses of 26,000 patient cases and 27,000 controls. *J. Clin. Oncol.*, **26**, 542–548.
 10. Park, D.J., Lesueur, F., Nguyen-Dumont, T., Pertesi, M., Odefrey, F., Hammet, F., Neuhausen, S.L., John, E.M., Andrulis, I.L., Terry, M.B. et al. (2012) Rare mutations in XRCC2 increase the risk of breast cancer. *Am. J. Hum. Genet.*, **90**, 734–739.
 11. Hilbers, F.S., Wijnen, J.T., Hoogerbrugge, N., Oosterwijk, J.C., Collee, M.J., Peterlongo, P., Radice, P., Manoukian, S., Feroce, I., Capra, F. et al. (2012) Rare variants in XRCC2 as breast cancer susceptibility alleles. *J. Med. Genet.*, **49**, 618–620.
 12. Thompson, E.R., Doyle, M.A., Ryland, G.L., Rowley, S.M., Choong, D.Y., Tothill, R.W., Thorne, H., Barnes, D.R., Li, J., Ellul, J. et al. (2012) Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles. *PLoS Genet.*, **8**, e1002894.
 13. Ellis, N.A. and Offit, K. (2012) Heterozygous mutations in DNA repair genes and hereditary breast cancer: a question of power. *PLoS Genet.*, **8**, e1003008.
 14. Bakker, J.L., van Mil, S.E., Crossan, G., Sabbaghian, N., De Leeener, K., Poppe, B., Adank, M., Gille, H., Verheul, H., Meijers-Heijboer, H. et al. (2013) Analysis of the novel fanconi anemia gene SLX4/FANCP in familial breast cancer cases. *Hum. Mutat.*, **34**, 70–73.
 15. de Garibay, G.R., Diaz, A., Gavina, B., Romero, A., Garre, P., Vega, A., Blanco, A., Tosar, A., Diez, O., Perez-Segura, P. et al. (2013) Low prevalence of SLX4 loss-of-function mutations in non-BRCA1/2 breast and/or ovarian cancer families. *Eur. J. Hum. Genet.*, **21**, 883–886.
 16. Shah, S., Kim, Y., Ostrovskaya, I., Murali, R., Schrader, K.A., Lach, F.P., Sarrel, K., Rau-Murthy, R., Hansen, N., Zhang, L. et al. (2013) Assessment of SLX4 mutations in hereditary breast cancers. *PLoS One*, **8**, e66961.
 17. Gracia-Aznarez, F.J., Fernandez, V., Pita, G., Peterlongo, P., Dominguez, O., de la Hoya, M., Duran, M., Osorio, A., Moreno, L., Gonzalez-Neira, A. et al. (2013) Whole exome sequencing suggests much of non-BRCA1/BRCA2 familial breast cancer is due to moderate and low penetrance susceptibility alleles. *PLoS One*, **8**, e55681.
 18. Wang, G.S. and Cooper, T.A. (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat. Rev. Genet.*, **8**, 749–761.
 19. Sterne-Weiler, T., Howard, J., Mort, M., Cooper, D.N. and Sanford, J.R. (2011) Loss of exon identity is a common mechanism of human inherited disease. *Genome Res.*, **21**, 1563–1571.
 20. Mucaki, E.J., Shirley, B.C. and Rogan, P.K. (2013) Prediction of mutant mRNA splice isoforms by information theory-based exon definition. *Hum. Mutat.*, **34**, 557–565.
 21. Caceres, J.F., Stamm, S., Helfman, D.M. and Krainer, A.R. (1994) Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science*, **265**, 1706–1709.
 22. Meetei, A.R., Medhurst, A.L., Ling, C., Xue, Y., Singh, T.R., Bier, P., Steltenpool, J., Stone, S., Dokal, I., Mathew, C.G. et al. (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat. Genet.*, **37**, 958–963.
 23. Kiiski, J.I., Pelttari, L.M., Khan, S., Freysteinsdottir, E.S., Reynisdottir, I., Hart, S.N., Shimelis, H., Vilske, S., Kallioniemi, A., Schleutker, J. et al. (2014) Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. *Proc. Natl Acad. Sci. USA*, **111**, 15172–15177.
 24. Kottemann, M.C. and Smogorzewska, A. (2013) Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature*, **493**, 356–363.
 25. Tischkowitz, M. and Xia, B. (2010) PALB2/FANCM: recombining cancer and Fanconi anemia. *Cancer Res.*, **70**, 7353–7359.
 26. Walden, H. and Deans, A.J. (2014) The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. *Annu. Rev. Biophys.*, **43**, 257–278.
 27. Huang, J., Liu, S., Bellani, M.A., Thazhathveetil, A.K., Ling, C., de Winter, J.P., Wang, Y., Wang, W. and Seidman, M.M. (2013) The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. *Mol. Cell*, **52**, 434–446.
 28. Antoniou, A.C., Casadei, S., Heikkinen, T., Barrowdale, D., Pylkas, K., Roberts, J., Lee, A., Subramanian, D., De Leeener, K., Fostira, F. et al. (2014) Breast-cancer risk in families with mutations in PALB2. *N. Engl. J. Med.*, **371**, 497–506.
 29. Sawyer, S.L., Tian, L., Kähkönen, M., Schwartzentruber, J. and Kircher, M.; University of Washington Centre for Mendelian Genomics; FORGE Canada Consortium, Majewski, J., Dymont, D.A., Innes, A.M. et al. (2015) Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. *Cancer Discov.*, **5**, 135–142.
 30. Singh, T.R., Bakker, S.T., Agarwal, S., Jansen, M., Grassman, E., Godthelp, B.C., Ali, A.M., Du, C.H., Rooimans, M.A., Fan, Q. et al. (2009) Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. *Blood*, **114**, 174–180.
 31. Lim, E.T., Würtz, P., Havulinna, A.S., Palta, P., Tukiainen, T., Rehnström, K., Esko, T., Mägi, R., Inouye, M. and Lappalainen, T. et al. (2014) Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet.*, **10**, e1004494.
 32. Southey, M.C., Park, D.J., Nguyen-Dumont, T., Campbell, I., Thompson, E., Trainer, A.H., Chenevix-Trench, G., Simard, J., Dumont, M., Soucy, P. et al. (2013) COMPLEXO: identifying the missing heritability of breast cancer via next generation collaboration. *Breast Cancer Res.*, **15**, 402.
 33. Hosmer, D.W. and Lemeshow, S. (1989) *Applied Logistic Regression*. John Wiley & Sons, New York, USA.
 34. Normand, S.L. (1999) Tutorial in biostatistics. Meta-analysis: formulating, evaluating, combining and reporting. *Stat. Med.*, **18**, 321–359.

35. Mehta, C.R., Patel, N. and Senchaudhuri, P. (2000) Efficient Monte Carlo methods for conditional logistic regression. *J. Am. Stat. Assoc.*, **95**, 99–108.
36. Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B. Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S. et al. (2012) Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep.*, **1**, 167–178.
37. Bi, C. and Rogan, P.K. (2004) Bipartite pattern discovery by entropy minimization-based multiple local alignment. *Nucleic Acids Res.*, **32**, 4979–4991.
38. Olsen, R.K., Broner, S., Sabaratnam, R., Doktor, T.K., Andersen, H.S., Bruun, G.H., Gahrn, B., Stenbroen, V., Olpin, S.E., Dobbie, A. et al. (2014) The ETFDH c.158A>G variation disrupts the balanced interplay of ESE- and ESS-binding proteins thereby causing missplicing and multiple Acyl-CoA dehydrogenation deficiency. *Hum. Mutat.*, **35**, 86–95.
39. Fackenthal, J.D., Cartegni, L., Krainer, A.R. and Olopade, O.I. (2002) BRCA2 T2722R is a deleterious allele that causes exon skipping. *Am. J. Hum. Genet.*, **71**, 625–631.
40. Bruun, G.H., Doktor, T.K. and Andresen, B.S. (2013) A synonymous polymorphic variation in ACADM exon 11 affects splicing efficiency and may affect fatty acid oxidation. *Mol. Genet. Metab.*, **110**, 122–128.
41. Pastor, T. and Pagani, F. (2011) Interaction of hnRNPA1/A2 and DAZAP1 with an Alu-derived intronic splicing enhancer regulates ATM aberrant splicing. *PLoS One*, **6**, e23349.
42. Colombo, M., De Vecchi, G., Caleca, L., Foglia, C., Ripamonti, C.B., Ficarazzi, F., Barile, M., Varesco, L., Peissel, B., Manoukian, S. et al. (2013) Comparative in vitro and in silico analyses of variants in splicing regions of BRCA1 and BRCA2 genes and characterization of novel pathogenic mutations. *PLoS One*, **8**, e57173.
43. Bakker, S.T., van de Vrugt, H.J., Rooimans, M.A., Oostra, A.B., Steltenpool, J., Delzenne-Goette, E., van der Wal, A., van der Valk, M., Joenje, H., te Riele, H. et al. (2009) Fanconi-deficient mice reveal unique features of Fanconi anemia complementation group M. *Hum. Mol. Genet.*, **18**, 3484–3495.
44. Bogliolo, M., Schuster, B., Stoepker, C., Derkunt, B., Su, Y., Raams, A., Trujillo, J.P., Minguillon, J., Ramirez, M.J., Pujol, R. et al. (2013) Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am. J. Hum. Genet.*, **92**, 800–806.
45. Castella, M., Pujol, R., Callen, E., Trujillo, J.P., Casado, J.A., Gille, H., Lach, F.P., Auerbach, A.D., Schindler, D., Benitez, J. et al. (2011) Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. *Blood*, **117**, 3759–3769.

Anexo 5

Bogliolo M, Surrallés J. Fanconi anemia: a model disease for studies on human genetics and advanced therapeutics. Curr Opin Genet Dev. 2015 Aug 6;33:32-40.

Fanconi anemia: a model disease for studies on human genetics and advanced therapeutics

Massimo Bogliolo² and Jordi Surrallés^{1,2}



Fanconi anemia (FA) is characterized by bone marrow failure, malformations, and chromosome fragility. We review the recent discovery of FA genes and efforts to develop genetic therapies for FA in the last five years. Because current data exclude FANCM as an FA gene, 15 genes remain bona fide FA genes and three (*FANCO*, *FANCR* and *FANCS*) cause an FA like syndrome. Monoallelic mutations in 6 FA associated genes (*FANCD1*, *FANCI*, *FANCF*, *FANCG*, *FANCL* and *FANCL*) predispose to breast and ovarian cancer. The products of all these genes are involved in the repair of stalled DNA replication forks by unhooking DNA interstrand cross-links and promoting homologous recombination. The genetic characterization of patients with FA is essential for developing therapies, including hematopoietic stem cell transplantation from a sibling donor after embryo selection, gene therapy, or genome editing using genetic recombination or engineered nucleases. Newly acquired knowledge about FA promises to provide therapeutic strategies in the near future.

Addresses

¹ Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

² Centre for Biomedical Network Research on Rare Diseases (CIBERER), Spain

Corresponding author: Surrallés, Jordi (jordi.surralles@uab.cat)

Current Opinion in Genetics & Development 2015, 33:32–40

This review comes from a themed issue on **Molecular and genetic bases of disease**

Edited by **Dan E Arking** and **Johanna M Rommens**

<http://dx.doi.org/10.1016/j.gde.2015.07.002>

0959-437/© 2015 Elsevier Ltd. All rights reserved.

Introduction: Fanconi anemia and Fanconi anemia-like genes

Fanconi anemia (FA), which affects approximately 1–3 of 500 000 newborns, causes bone marrow failure (BMF), malformations, and cancer predisposition. Hallmarks of FA are chromosome fragility and hypersensitivity to drugs that induce DNA interstrand cross-links (ICLs). Numerous other physiological and cellular abnormalities likely contribute to pathogenesis (Figure 1). The current decade has been prolific for the discovery of novel FA genes.

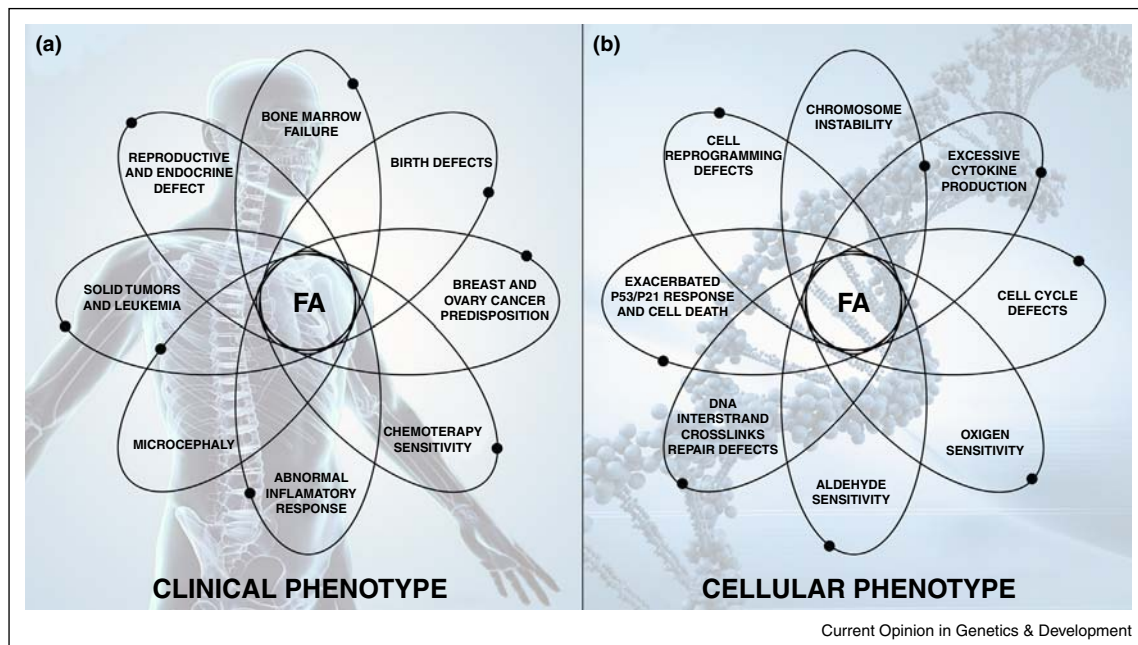
Thus, 19 genes are associated with FA, including the recently discovered genes *FANCO/RAD51C* [1^{**}], *FANCP/SLX4* [2^{**},3^{**}], *FANCO/ERCC4* [4^{**}], *FANCR/RAD51* [5^{**}], *FANCS/BRCA1* [6^{**}] and *FANCI/UBE2T* [7^{**},8^{**},9^{**}]. However, following a stringent criteria of at least two patients with BMF and a positive chromosome fragility test, only 15 are classified as bona fide FA genes (*FANCA*, *B*, *C*, *D1*, *D2*, *E*, *F*, *G*, *I*, *J*, *L*, *N*, *P*, *Q* and *T*). *FANCO*, *FANCR* and *FANCS* are FA-like genes, because they cause a chromosome fragility syndrome with FA-related malformations but without BMF (Figure 2). *FANCM* should also be excluded from the list of FA genes, because only one patient is reported to carry biallelic mutations [10]. It was subsequently found that this individual also carries biallelic *FANCA* mutations and has a sibling subtyped as FANCA. Notably, FANCM fails to complement the genetic defect of this patient [11]. Moreover, loss-of-function FANCM variants are more common than originally predicted and two individuals of Finnish descent with loss of function *FANCM* variants are healthy and exhibit normal hematology [12^{*}]. Thus, we recommend excluding *FANCM* as an FA gene, although, together with FAAP100, FAAP25, and other FA-core complex interacting proteins, FANCM is involved in the FA ICL repair (ICLR) pathway (see below). Similarly, whole exome sequencing (WES) detected biallelic *XRCC2* mutations in a consanguineous FA family [13]; however, this is again a single case and, because of the lack of genetic complementation data or any other functional evidence of a causative role of this homozygous mutation in disease, *XRCC2* should not yet be considered an FA gene.

The genetic heterogeneity and the number of private and founder mutations make the mutational analysis of FA patients extremely difficult and time consuming by traditional techniques such as Sanger sequencing [14–16]. However, next-generation sequencing (NGS) technologies, including WES or targeted sequencing of FA genes, together with high-resolution methods to detect large deletions, such as comparative genome hybridization arrays, single-nucleotide polymorphism arrays, and targeted Multiplex Ligation-dependent Probe Amplification, are now being implemented to find the underlying FA gene and determine the pathogenic mutations of new patients with FA [4^{**},17–21].

FA genes that predispose to breast and ovarian cancer

FANCD1/BRCA2, *FANCS/BRCA1*, *FANCI/BRIP1*, *FANCM*, *FANCN/PALB2*, and *FANCO/RAD51C* are breast

Figure 1



FA clinical and cellular phenotypes. **(a)** A wide range of clinical features, all of them of incomplete penetrance, characterizes the FA clinical phenotype. The main FA clinical features can be organized in four fundamental categories: abnormal embryo development (birth defects, microcephaly), bone marrow failure, reproductive and endocrine defects and cancer predisposition including solid tumors and leukemia. In some genetic subtypes the disease is linked to breast and ovarian cancer predisposition in monoallelic carriers and some adult FA patients. **(b)** The hallmark of FA cells is a DNA repair defect that causes cellular sensitivity to ICL-inducing agents, chromosome instability and cell cycle alterations. FA cells have several other phenotypic abnormalities most probably related to the DNA repair defect such as the oxygen and aldehyde sensitivity, increased cell death and the cell reprogramming defects. Additionally, FA cells overproduce proinflammatory cytokines that are known to be proapoptotic for HSC. All these characteristics are probably related to each other and their combination together with stochastic factors account for the clinical phenotype of FA patients.

and ovarian cancer susceptibility genes in carriers of monoallelic mutations (Figure 1) highlighting the fundamental link between FA and familial breast and ovarian cancer (FBOC). *RAD51 C* mutations influence ovarian cancers more than breast cancers [22,23], and are linked to other tumors such as head and neck cancer [24,25]. *RAD51 C* and *FANCM* were initially associated to FA before they were candidates for FBOC in monoallelic carriers [26,27] highlighting the fundamental role of FA research in advancing molecular oncology.

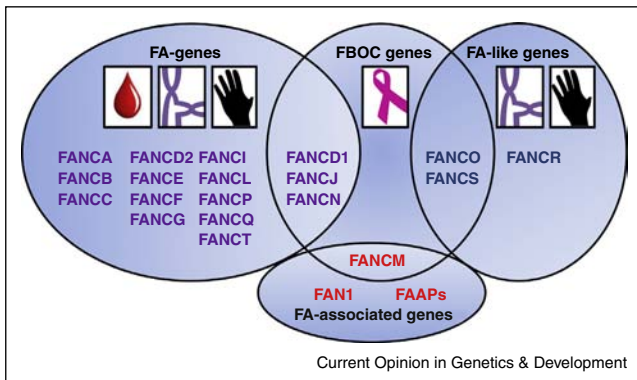
Historically, only FA genes (see below; Figure 2) with a direct role in the homologous recombination repair (HRR) late step of the FA pathway are linked to FBOC [28,29]. The two recently identified FA genes *FANCP/SLX4* and *FANCO/ERCC4* were also excluded as major breast cancer susceptibility genes in Italian, German, Spanish, Estonian, Jewish, and non-Jewish American populations [30–37]. However, pathogenic mutations in *FANCM* are associated with breast cancer susceptibility in several populations [27,38], suggesting a core complex-independent role (see below) of *FANCM*. In fact, even though *FANCM* is not essential for *RAD51* foci formation and HRR, the camptothecin sensitivity of *FANCM*

cells is shared with *FANCD1* and *FANCN* cells, linking *FANCM* to the step of the FA pathway connected to HRR [11].

The Fanconi anemia pathway: find, unhook, bypass, and recombine

ICLs are highly damaging, because they impede transcription and replication-fork progression. Since they affect both DNA strands, ICLs complicate error-free DNA repair, because an undamaged DNA template is not available. The FA DNA repair pathway coordinates reactions that remove ICL damage to restore genome integrity (for more detailed reviews, see [39,40]). Eight FA proteins (*FANCA*, *FANCB*, *FANCC*, *FANCE*, *FANCF*, *FANCG*, *FANCL*, and *FANCM*) form a nuclear complex (FANCore) whose ubiquitin E3 ligase function is activated by blocked DNA replication forks (Figure 3). The activated FANCore complex monoubiquitinates the *FANCD2*–*FANCI* heterodimer (ID complex) and *UBE2T/FANCT* provides the E2 conjugase activity necessary to this process [41–43]. The activated ID complex relocates to the damaged DNA in an ATR and *BRCA1*-dependent manner [41,43,44], promotes nucleolytic cleavage of the 3' and 5' sites of DNA to

Figure 2

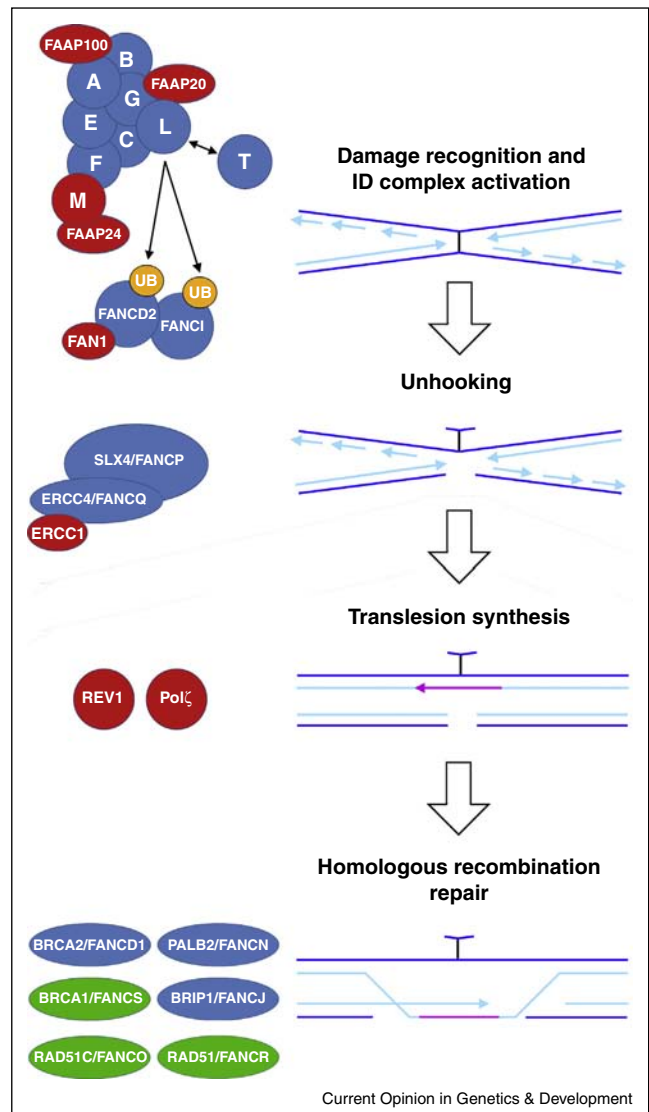


FA related genes. Biallelic mutation in 15 genes (*FANCA, B, C, D1/BRCA2, D2, E, F, G, I, J/BRIP1, L, N/PALB2, P/SLX4, Q/ERCC4* and *T/UBE2T*; marked in purple) cause FA, characterized by the triad of bone marrow failure, chromosome fragility and malformations. Biallelic mutations in three genes (*FANCO/RAD51 C, FANCR/RAD51* and *FANCS/BRIP1*; marked in red) cause a FA-like syndrome without bone marrow failure. There are a number of FA-associated genes not linked to a FA disease phenotype in biallelic mutation carriers including *FAAP20 (C1orf86), FAAP24 (C19orf40), FAAP100 (C17orf70)* and *FAAP250 (FANCM)*. *FANCM* is therefore excluded from the list of FA genes. Six of the above genes (*FANCD1, J, M, N, O, and S*; marked in dark blue) are associated to familiar breast and ovarian cancer in monoallelic mutation carriers highlighting the fundamental link between FA and cancer predisposition in the general population.

unhook the ICL and, successively, induces trans-lesion polymerases REV1 and POL ζ [45,46,47^{**},48]. These reactions extend the leading DNA strand above and past the unhooked ICL to produce a substrate that is processed by successive HRR reactions [39].

In the last few years, major advances in understanding the unhooking step of ICL repair came from the discovery of the FA genes *SLX4/FANCP* and *ERCC4/FANCO*. Current models predict that DNA integrity is not restored without cleaving the 3' and 5' sequences flanking the lesion. Several nucleases contribute to ICLR, such as XPF-ERCC1, MUS81-EME1, SLX1, SNM1A, and SNM1B [49–51]. However, the lack of FA patients with mutations in the genes encoding these nucleases hindered the identification of the main FA/BRCA pathway nuclease. BTBD12/SLX4 provides a platform for several endonucleases involved in ICL repair, including ERCC4-ERCC1, MUS81-EME1, and SLX1 that dock to cleave DNA flaps, replication forks, and Holliday junctions [52–54]. *SLX4* is mutated in patients with bona fide FA (*FANCP*) [2^{**},3^{**}]. In contrast to other FA mouse models, the *Slx4* knockout (KO) mouse exhibits an FA phenotype with developmental defects and cytopenia [55^{**}]. The N-terminal segment of *SLX4/FANCP* harbors ubiquitin zinc finger (UBZ) domains, suggesting that *SLX4* is recruited to DNA damage via interaction with ubiquitinated proteins involved in the DNA damage response

Figure 3



The FA/BRCA DNA ICL repair pathway: FA core complex is activated upon stalled replication forks and monoubiquitinates the ID complex. Activated ID complex relocates to chromatin and promotes SLX4 and ERCC4 activities in ICL unhooking and subsequent translesion synthesis. In the final step of the repair pathway, genome integrity is restored by homologous recombination repair upon the action of downstream FA proteins. Proteins coded by bona fide FA genes are shown in blue while the products of FA-like genes are shown in green. The FA associated proteins whose genes are not mutated in any FA or FA-like patient are shown in red. FA Core complex (FANCore) proteins are referred to by designated characters only instead of complete names in order to fit them all in a smaller space.

(DDR). *SLX4* recruitment in chicken DT-40 cells may depend on FANCD2 monoubiquitination but with uncertain relevance to mammals [56]. *SLX4* serves as a scaffold that organizes specific nucleases for transport to DNA lesions and regulates nuclease activity [56,57]. *SLX4* contains SUMO-interacting motifs (SIMs)

required for binding sumoylated DNA repair proteins, and SLX4 acts directly or indirectly as a SUMO E3 ligase. These SUMO-related functions are not required for ICL repair but for a general response to replication stress. Accordingly, mutations of SLX4 SIMs do not produce ICLs hypersensitivity but cause common fragile site instability and increased mitotic catastrophe [58,59].

ERCC4 mutations were identified using WES and Sanger sequencing in two unrelated and unassigned FA patients [4**]. Both patients had characteristic FA symptoms, including BMF, chromosome fragility, and FA-related malformations; *ERCC4* was therefore renamed *FANCO* [4**]. ERCC4–ERCC1 is a heterodimeric endonuclease originally described as an essential component of the nucleotide excision repair (NER) system. The gene coding the catalytic subunit ERCC4, also known as XPF, is mutated in patients with Xeroderma Pigmentosum (XP), complementation group F [60]. During NER, ERCC4/ERCC1 endonuclease is responsible for the incision on the 5' side of DNA lesions but, in contrast to other XP-related proteins, defects in XPF sensitize cells to both UV light and ICLs, indicating an NER-independent DNA repair role [61]. NER and ICLR functions of ERCC4/ERCC1 can be artificially uncoupled by mutating ERCC1 residues important for the interaction of ERCC4/ERCC1 with XPA. Cells expressing this mutant form of ERCC1 are sensitive to UVC irradiation but not to ICLs [62]. Similarly, FA causing mutations in ERCC4 naturally uncouple ERCC4/*FANCO* NER and ICLR functions with inhibition of ICLR but not NER activity. In fact FA-Q patients lacked XP symptoms, consistent with functional NER despite ERCC4/*FANCO* defects [4**].

The Fanconi anemia-associated nuclease 1 (FAN1) was identified in 2010 and immediately stole the scene as the main candidate FA nuclease [63–66]. FAN1 is a structure-specific nuclease that, when depleted, causes sensitivity to ICLs. Moreover, FAN1 is recruited to damaged DNA via the interaction through its N-terminal UBZ domain with monoubiquitinated FANCD2. Recruitment of FAN1 through FANCD2 monoubiquitination is consistent with the loss of nucleolytic incisions near the ICL *in vitro* upon FANCD2 depletion [47**]. Although FAN1 interacts with FANCD2 and contributes to the FA/BRCA pathway of ICLR [67–70], FAN1 deficiency is not associated with an FA phenotype [71]. Instead, biallelic loss-of-function point mutations in *FAN1* cause karyomegalic interstitial nephritis, linking chronic kidney failure to ICLR [72] and monoallelic *FAN1* mutations are linked to familial colon cancer [73]. Further, in contrast to ERCC4–ERCC1, FAN1 activity (similarly to MUS81–EME1 activity) is not required for nucleolytic incisions near an ICL [74]. Therefore, the role of FAN1 nuclease activity in the FA pathway is still under discussion.

Fanconi anemia and endogenous aldehydes

Groundbreaking series of studies led by Dr KJ Patel's team has shed light on understanding the endogenous DNA damage that causes FA as they demonstrated using genetic models that the FA DNA repair pathway counteracts the genotoxicity of endogenous aldehydes [75**]. Aldh2, which mediates the metabolism of acetaldehyde, is essential for embryonic development of FA mice and dual deficiency of ALDH5 and the FA pathway mediate synthetic lethality in DT40 chicken cells [75**,76]. *Aldh2*^{-/-} *Fancd2*^{-/-} KO mice are not viable if the mother lacks functional Aldh2. When the mother is *Aldh2*^{+/-}, *Aldh2*^{-/-} *Fancd2*^{-/-} offspring are viable, demonstrating that maternal aldehyde catabolism rescues embryo lethality. HSCs of viable *Aldh2*^{-/-} *Fancd2*^{-/-} mice spontaneously suffer increased DNA damage, and the mice develop acute leukemia and aplastic anemia. Double KO mice have a 30-fold reduction in the HSC pool when compared to wild type mice and a 10-fold reduction when compared to singly deficient *Fancd2*^{-/-} mice. Further, there is a 600-fold reduction in the bone marrow reconstitution ability of HSC from mice deficient in the FA DNA repair pathway and the detoxification of acetaldehyde when compared to wild type mice [77*]. When similar experiments are performed by knocking out *Fanca* (which is mutated in the majority of FA patients) together with *Aldh2*, the resulting phenotype is even worse. *Aldh2*^{-/-} *Fanca*^{-/-} embryos are mostly resorbed around stage E10–E12.5 of development if mothers are *Aldh2*^{+/-}. If implanted in *Aldh2*^{+/-} mothers, double KO embryos can develop after stage E12.5 but they present severe malformations, affecting especially limbs and eyes, and they die before birth. Only when *Aldh2*^{-/-} *Fanca*^{-/-} embryos are transferred to *Aldh2*^{+/+} mothers do viable offspring result, but these neonates have low numbers of hematopoietic stem and progenitor cells, indicating that fetal Aldh2 is essential for proper hematopoiesis [78]. These findings are relevant to Asian populations, particularly that of Japan with a 40% carrier frequency of a dominant-negative *ALDH2* allele. In Japanese patients with FA that bear the dominant-negative *ALDH2* allele, the FA phenotype is more severe with earlier onset of BMF and increased FA-related birth defects [79], which provides strong evidence that mouse data informs human disease. It remains to be determined whether targeting aldehyde metabolism ameliorates BMF and cancer predisposition of patients with FA [80–82]. *ALDH2*-deficient Japanese who consume alcohol are at higher risk of developing esophageal cancer, macrocytosis, and macrocytic anemia and oral microorganisms produce high levels of acetaldehyde in saliva [83]. These data suggest ways to prevent BMF as well as head and neck cancer in FA patients through reduction of alcohol consumption, improved oral hygiene and, possibly, by aldehyde chelation.

Novel therapies: from genes to patients

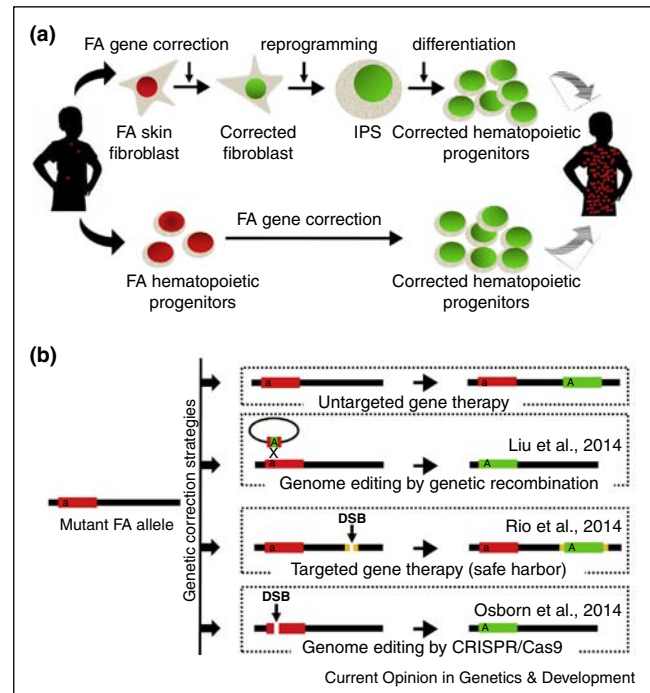
The discovery of innovative therapies highlights the pioneering role of FA translational research in the history

of medicine. Eliane Gluckman, one of the leading hematologists worldwide, performed the first umbilical cord-blood (UCB) transplant in Paris to treat a patient with FA [84]. Cord blood banking is now performed worldwide, and more than 30 000 patients with blood disorders and other diseases benefited from this source of blood progenitor cells [85]. The outcome of transplanting patients with FA using HLA-matched unrelated HSC donors doubled because of drugs such as fludarabine and improved protocols, approaching the excellent survival rates using HLA-matched donor siblings [86,87]. The first preimplantation genetic diagnosis combined with HLA-matching, generated a savior baby to treat a sibling with FA using UCB transplantation [88]. Hundreds of children with numerous blood disorders were subsequently treated [89,90]. However, FA families should be informed of the low success rate of this procedure (<5% of babies born per *in vitro* fertilization cycle) due to Mendelian restrictions and high aneuploidy rates associated with advanced maternal age [91].

Unfortunately, HSC donors are not available for all patients with FA, and HSCT, even with gentle myeloablative methods, are associated with cancer risk. To overcome these limitations, FA gene therapy clinical trials are in progress [92]. Difficulties in collecting sufficient blood progenitor cells from patients with FA and inefficient transduction protocols with first-generation retroviral vectors led to unsuccessful clinical trials a decade ago [93–95]. Improvements may come from a safer and more efficient lentiviral vector expressing human FANCA using the weak PGK promoter [96,97] developed by Bueren's laboratory, which was designated orphan drug by the European Medicines Agency. Drugs such as plerixafor efficiently mobilize HSCs for apheresis collection and this approach can be useful in patients with FA. Studies of mosaic patients and mice with FA after *ex vivo* gene therapy indicate a survival advantage of genetically corrected cells *in vivo*, even in the absence of myeloablative conditioning regimens [98,99]. These measures, together with the expected lack of graft-versus-host disease after gene therapy, may be a future therapeutic option to treat BMF of patients with FA and prevent HSCT-related cancers.

FA research also played a pioneering role in the field of regenerative medicine (Figure 4). Disease-free blood progenitor cells were first generated from the skin of a patient with FA via induced pluripotent stem (IPS) cells [100**]. Ultimately, sufficient IPS cell-derived HSCs must be generated for autotransplantation. Because FA fibroblasts are difficult to reprogram, this study uncovered a novel role of the FA pathway in cell reprogramming. Reprogramming induces the DDR and activates the FA pathway, leading to P53-mediated apoptosis and low reprogramming efficiency. Consequently, correcting FA genes restores the reprogramming efficiency of FA

Figure 4



Potential cell and gene therapy strategies in Fanconi anemia. (a) Gene corrected blood progenitors can be generated after genetic correction of patient-derived hematopoietic stem cells. This is classically known as *ex vivo* gene therapy. Alternatively, blood progenitors can be generated from disease-free IPS cells after reprogramming gene-corrected skin fibroblasts. Corrected blood progenitors would finally be used to treat the disease by autotransplantation. (b) There are several strategies for FA gene correction. Ongoing FA gene therapy clinical trials are based on untargeted gene therapy using lentiviral vectors carrying a wild type copy of the FANCA gene. Novel genetic tools allow targeted correction of the endogenous FANCA mutation by homologous recombination by the use of either helper dependent adenoviral vectors (HDAV) or engineered nucleases such as CRISPR/Cas9. Alternatively, genetic correction may be performed by targeted insertion of the wild type FANCA gene into a safe harbor locus with the help of engineered nucleases such as TALEN.

fibroblasts to IPS cells [100**,101,102,103*,104–106]. This is partially circumvented by preventing reactive oxygen species (ROS)-mediated DNA damage [101] or by suppressing P53 during reprogramming using RNA interference [103*] or human papillomavirus P53-repressing E6 protein [107]. Safe and controlled FANCA gene correction was achieved using integration-free genome editing by genetic recombination with helper-dependent adenoviral vectors [103*] or by targeting FANCA insertion into the safe locus AAVS1 [104] using engineered nucleases ('safe harbor' strategy [108]). Genome editing using CRISPR/Cas9-engineered nucleases corrects FANCA mutations in human FA fibroblasts [109*].

Although clinical translation of gene-corrected IPS cells and genome editing with engineered nucleases is difficult

in the short term, successful engraftment of iPSC-derived and gene-corrected blood progenitor cells may provide a future therapeutic approach for FA and other blood diseases characterized by low numbers of bone marrow HSCs [104,108]. Therefore, editing fibroblasts and IPS-cell genomes may soon translate directly to HSCs for FA patients (Figure 4a). FA-IPS cells offer a novel tool to model FA physiology and pathogenesis and provide a cell platform for drug screening [103*]. Finally, therapies designed to enhance the correct mRNA processing at a mutant TT splice donor in FANCC using suppressor U1 snRNAs, suggests that correcting pathological mRNA processing at specific mutant splice sites might apply to FA complementation groups in a mutation-specific fashion [110]. Therefore, our better understanding of the molecular genetic defects of FA and the use of gene correction strategies may contribute to futures treatments for this devastating disease.

Conflicts interest

The authors declare no conflicts of interest.

Acknowledgements

This review article is dedicated in remembrance of Prof. Johan de Winter (VUMC-Amsterdam) whose contribution to the genetics of Fanconi anemia will last forever. We would like to thank Dr Agata Smogorzewska (Rockefeller University, New York) for sharing unpublished data. Surrallés' laboratory is currently funded by the Generalitat de Catalunya (SGR0489-2009 and SGR317-2014), the ICREA-Academia Program, the Marató de TV3 (project 464/C/2012), the Spanish Ministry of Science and Innovation (projects CB06/07/0023 and SAF2012-31881), the European Commission (EUROFANCOLEN project HEALTH-F5-2012-305421) and the European Regional Development FEDER Funds. CIBERER is an initiative of the Instituto de Salud Carlos III, Spain.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Vaz F, Hanenberg H, Schuster B, Barker K, Wiek C, Erven V, •• Neveling K, Endt D, Kesterton I, Autore F *et al.*: **Mutation of the RAD51C gene in a Fanconi anemia-like disorder.** *Nat Genet* 2010, **42**:406-409.

This paper first describes homozygous missense mutations in the RAD51C gene in a consanguineous family with an FA-like phenotype.

2. Stoepker C, Hain K, Schuster B, Hilhorst-Hofstee Y, •• Rooimans MA, Steltenpool J, Oostra AB, Eirich K, Korthof ET, Nieuwint AW *et al.*: **SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype.** *Nat Genet* 2011, **43**:138-141.

This article along with Ref [3**] independently described several FA individuals with pathogenic biallelic mutations in SLX4.

3. Kim Y, Lach FP, Desetty R, Hanenberg H, Auerbach AD, •• Smogorzewska A: **Mutations of the SLX4 gene in Fanconi anemia.** *Nat Genet* 2011, **43**:142-146.

See annotation to Ref [2**].

4. Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A, •• Trujillo JP, Minguillon J, Ramirez MJ, Pujol R *et al.*: **Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia.** *Am J Hum Genet* 2013, **92**:800-806.

This study shows that a subset of mutations in ERCC4 specifically disrupts the function of XPF in ICLR causing FA subtype Q.

5. Wang ATT, Kim JE, Wagner B, Conti FP, Lach A, Huang H, •• Molina E, Sanborn H, Zierhut BK, Cornes A, Abhyankar SC,

Gabriel S, Auerbach AD, Kowalczykowski SC, Smogorzewska A: **A dominant mutation in human RAD51 reveals its function in DNA interstrand crosslink repair independent of homologous recombination.** *Mol Cell* 2015. (in press).

This paper along with Ref [6**] described that a dominant *de novo* mutation in FANCR and biallelic mutations in BRCA1, respectively, cause an FA-like phenotype with chromosome fragility but without BMF.

6. Sawyer SL, Tian L, Kahkonen M, Schwartzentruber J, Kircher M, •• Majewski J, Dymont DA, Innes AM, Boycott KM, Moreau LA *et al.*: **Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype.** *Cancer Discov* 2015, **5**(2):135-142.

See annotation to Ref [5**].

7. Hira A, Yoshida K, Sato K, Okuno Y, Shiraishi Y, Chiba K, •• Tanaka H, Miyano S, Shimamoto A, Tahara H *et al.*: **Mutations in the gene encoding the E2 conjugating enzyme UBE2T cause Fanconi anemia.** *Am J Hum Genet* 2015, **96**:1001-1007.

This reference along with [8**,9**] describe that biallelic mutations in the E2 conjugase UBE2T cause FA in several unrelated individuals.

8. Rickman KA, Lach FP, Abhyankar A, Donovan FX, Sanborn EM, •• Kennedy JA, Sougnez C, Gabriel SB, Elemento O, Chandrasekharappa SC *et al.*: **Deficiency of UBE2T, the E2 ubiquitin ligase necessary for FANCD2 and FANCI ubiquitination, causes FA-T subtype of Fanconi anemia.** *Cell Rep* 2015, **12**:35-41.

See annotation to Ref [7**].

9. Virts EL, Jankowska A, Mackay C, Glaas MF, Wiek C, Kelich SL, •• Lottmann N, Kennedy FM, Marchal C, Lehnert E *et al.*: **AluY-mediated germline deletion, duplication and somatic stem cell reversion in UBE2T defines a new subtype of Fanconi anemia.** *Hum Mol Genet* 2015. (Epub ahead of print).

See annotation to Ref [7**].

10. Meetei AR, Medhurst AL, Ling C, Xue Y, Singh TR, Bier P, Steltenpool J, Stone S, Dokal I, Mathew CG *et al.*: **A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M.** *Nat Genet* 2005, **37**:958-963.

11. Singh TR, Bakker ST, Agarwal S, Jansen M, Grassman E, Godthelp BC, Ali AM, Du CH, Rooimans MA, Fan Q *et al.*: **Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M.** *Blood* 2009, **114**:174-180.

12. Lim ET, Wurtz P, Havulinna AS, Palta P, Tukiainen T, Rehnstrom K, Esko T, Magi R, Inouye M, Lappalainen T *et al.*: **Distribution and medical impact of loss-of-function variants in the Finnish founder population.** *PLoS Genet* 2014, **10**:e1004494.

This paper described individuals with biallelic loss-of-function mutations in FANCM without an FA phenotype, definitively excluding FANCM as an FA causing gene.

13. Shamseldin HE, Elfaki M, Alkuraya FS: **Exome sequencing reveals a novel Fanconi group defined by XRCC2 mutation.** *J Med Genet* 2012, **49**:184-186.

14. Castella M, Pujol R, Callen E, Trujillo JP, Casado JA, Gille H, Lach FP, Auerbach AD, Schindler D, Benitez J *et al.*: **Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations.** *Blood* 2011, **117**:3759-3769.

15. Schuster B, Knies K, Stoepker C, Velleuer E, Friedl R, Gottwald-Muhlhauser B, de Winter JP, Schindler D: **Whole exome sequencing reveals uncommon mutations in the recently identified Fanconi anemia gene SLX4/FANCP.** *Hum Mutat* 2013, **34**(1):93-96.

16. Gille JJ, Floor K, Kerkhoven L, Ameziane N, Joenje H, de Winter JP: **Diagnosis of Fanconi anemia: mutation analysis by multiplex ligation-dependent probe amplification and PCR-based Sanger sequencing.** *Anemia* 2012, **2012**:603253.

17. Ameziane N, Sie D, Dentre S, Ariyurek Y, Kerkhoven L, Joenje H, Dorsman JC, Ylstra B, Gille JJ, Sierstern EA *et al.*: **Diagnosis of Fanconi anemia: mutation analysis by next-generation sequencing.** *Anemia* 2012, **2012**:132856.

18. Knies K, Schuster B, Ameziane N, Rooimans M, Bettecken T, de Winter J, Schindler D: **Genotyping of Fanconi anemia patients by whole exome sequencing: advantages and challenges.** *PLOS ONE* 2012, **7**:e52648.

19. Schuster B, Knies K, Stoepker C, Velleuer E, Friedl R, Gottwald-Mühlhauser B, de Winter JP, Schindler D: **Whole exome sequencing reveals uncommon mutations in the recently identified Fanconi anemia gene SLX4/FANCP.** *Hum Mutat* 2013, **34**:93-96.
20. Chandrasekharappa SC, Lach FP, Kimble DC, Kamat A, Teer JK, Donovan FX, Flynn E, Sen SK, Thongthip S, Sanborn E *et al.*: **Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia.** *Blood* 2013, **121**:e138-e148.
21. Flynn EK, Kamat A, Lach FP, Donovan FX, Kimble DC, Narisu N, Sanborn E, Boulad F, Davies SM, Gillio AP 3rd *et al.*: **Comprehensive analysis of pathogenic deletion variants in Fanconi anemia genes.** *Hum Mutat* 2014, **35**:1342-1353.
22. Loveday C, Turnbull C, Ruark E, Xicola RM, Ramsay E, Hughes D, Warren-Perry M, Snape K, Eccles D, Evans DG *et al.*: **Germine RAD51C mutations confer susceptibility to ovarian cancer.** *Nat Genet* 2012, **44**:475-476.
23. Osorio A, Endt D, Fernandez F, Eirich K, de la Hoya M, Schmutzler R, Caldes T, Meindl A, Schindler D, Benitez J: **Predominance of pathogenic missense variants in the RAD51C gene occurring in breast and ovarian cancer families.** *Hum Mol Genet* 2012, **21**:2889-2898.
24. Scheckenbach K, Baldus SE, Balz V, Freund M, Pakropa P, Sproll C, Schafer KL, Wagenmann M, Schipper J, Hanenberg H: **RAD51C — a new human cancer susceptibility gene for sporadic squamous cell carcinoma of the head and neck (HNSCC).** *Oral Oncol* 2014, **50**:196-199.
25. Gresner P, Gromadzinska J, Twardowska E, Rydzynski K, Wasowicz W: **Rad51C: a novel suppressor gene modulates the risk of head and neck cancer.** *Mutat Res Fundam Mol Mech Mutagen* 2014, **762**:47-54.
26. Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, Freund M, Lichtner P, Hartmann L, Schaal H *et al.*: **Germine mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene.** *Nat Genet* 2010, **42**:410-414.
- Once implicated in FA, RAD51C was described in this study to cause breast and ovarian cancer predisposition in monoallelic mutation carriers.
27. Kiiski JI, Pelttari LM, Khan S, Freysteinsdottir ES, Reynisdottir I, Hart SN, Shimelis H, Vilske S, Kallioniemi A, Schleutker J *et al.*: **Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer.** *Proc Natl Acad Sci U S A* 2014, **111**:15172-15177.
- This paper implicates FANCM in breast cancer predisposition.
28. Levy-Lahad E: **Fanconi anemia and breast cancer susceptibility meet again.** *Nat Genet* 2010, **42**:368-369.
29. D'Andrea AD: **Susceptibility pathways in Fanconi's anemia and breast cancer.** *N Engl J Med* 2010, **362**:1909-1919.
30. Landwehr R, Bogdanova NV, Antonenkova N, Meyer A, Bremer M, Park-Simon TW, Hillemanns P, Karstens JH, Schindler D, Dork T: **Mutation analysis of the SLX4/FANCP gene in hereditary breast cancer.** *Breast Cancer Res Treat* 2011, **130**:1021-1028.
31. Catucci I, Colombo M, Verderio P, Bernard L, Ficarazzi F, Mariette F, Barile M, Peissel B, Cattaneo E, Manoukian S *et al.*: **Sequencing analysis of SLX4/FANCP gene in Italian familial breast cancer cases.** *PLOS ONE* 2012, **7**:e31038.
32. Fernandez-Rodriguez J, Quiles F, Blanco I, Teule A, Feliubadalo L, Valle JD, Salinas M, Izquierdo A, Darder E, Schindler D *et al.*: **Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families.** *BMC Cancer* 2012, **12**:84.
33. Bakker JL, van Mil SE, Crossan G, Sabbaghian N, De Leeneer K, Poppe B, Adank M, Gille H, Verheul H, Meijers-Heijboer H *et al.*: **Analysis of the novel Fanconi anemia gene SLX4/FANCP in familial breast cancer cases.** *Hum Mutat* 2013, **34**:70-73.
34. de Garibay GR, Diaz A, Gavina B, Romero A, Garre P, Vega A, Blanco A, Tosar A, Diez O, Perez-Segura P *et al.*: **Low prevalence of SLX4 loss-of-function mutations in non-BRCA1/2 breast and/or ovarian cancer families.** *Eur J Hum Genet* 2013, **21**:883-886.
35. Shah S, Kim Y, Ostrovskaya I, Murali R, Schrader KA, Lach FP, Sarrel K, Rau-Murthy R, Hansen N, Zhang L *et al.*: **Assessment of mutations in hereditary breast cancers.** *PLOS ONE* 2013, **8**:e66961.
36. Osorio A, Bogliolo M, Fernández V, Barroso A, de la Hoya M, Caldes T, Lasa A, Cajal TRY, Santamariña M, Vega A *et al.*: **Evaluation of rare variants in the new Fanconi anemia gene ERCC4 (FANCC) as familial breast/ovarian cancer susceptibility alleles.** *Hum Mutat* 2013, **34**(12):1615-1618.
37. Kohlhase S, Bogdanova NV, Schurmann P, Bermisheva M, Khusnutdinova E, Antonenkova N, Park-Simon TW, Hillemanns P, Meyer A, Christiansen H *et al.*: **Mutation analysis of the ERCC4/FANCC gene in hereditary breast cancer.** *PLOS ONE* 2014, **9**:e85334.
38. Peterlongo P, Catucci I, Colombo M, Caleca L, Mucaki E, Bogliolo M, Marin M, Damiola F, Bernard L, Pensotti V *et al.*: **FANCM c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity, and is a familial breast cancer risk factor.** *Hum Mol Genet* 2015. (Epub ahead of print).
39. Walden H, Deans AJ: **The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder.** *Annu Rev Biophys* 2014, **43**:257-278.
40. Kottemann MC, Smogorzewska A: **Fanconi anaemia and the repair of Watson and Crick DNA crosslinks.** *Nature* 2013, **493**:356-363.
41. Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, Grompe M, D'Andrea AD: **Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway.** *Mol Cell* 2001, **7**:249-262.
42. Machida YJ, Machida Y, Chen Y, Gurtan AM, Kupfer GM, D'Andrea AD, Dutta A: **UBE2T is the E2 in the Fanconi anemia pathway and undergoes negative autoregulation.** *Mol Cell* 2006, **23**:589-596.
43. Smogorzewska A, Matsuoka S, Vinciguerra P, McDonald ER 3rd, Hurov KE, Luo J, Ballif BA, Gygi SP, Hofmann K, D'Andrea AD *et al.*: **Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair.** *Cell* 2007, **129**:289-301.
44. Bogliolo M, Lyakhovich A, Callen E, Castella M, Cappelli E, Ramirez MJ, Creus A, Marcos R, Kalb R, Neveling K *et al.*: **Histone H2AX and Fanconi anemia FANCD2 function in the same pathway to maintain chromosome stability.** *EMBO J* 2007, **26**:1340-1351.
45. Niedzwiedz W, Mosedale G, Johnson M, Ong CY, Pace P, Patel KJ: **The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair.** *Mol Cell* 2004, **15**:607-620.
46. Raschle M, Knipscheer P, Enoiu M, Angelov T, Sun J, Griffith JD, Ellenberger TE, Schärer OD, Walter JC: **Mechanism of replication-coupled DNA interstrand crosslink repair.** *Cell* 2008, **134**:969-980.
47. Knipscheer P, Raschle M, Smogorzewska A, Enoiu M, Ho TV, Schärer OD, Elledge SJ, Walter JC: **The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair.** *Science* 2009, **326**:1698-1701.
- This paper presents a forks collision model at single nucleotide resolution of ICLR pathway.
48. Kim H, Yang K, Dejsuphong D, D'Andrea AD: **Regulation of Rev1 by the Fanconi anemia core complex.** *Nat Struct Mol Biol* 2012, **19**:164-170.
49. Niedernhofer LJ, Lalai AS, Hoeijmakers JH: **Fanconi anemia (cross)linked to DNA repair.** *Cell* 2005, **123**:1191-1198.
50. Deans AJ, West SC: **DNA interstrand crosslink repair and cancer.** *Nat Rev Cancer* 2011, **11**:467-480.
51. Cattell E, Sengerova B, McHugh PJ: **The SNM1/Pso2 family of ICL repair nucleases: from yeast to man.** *Environ Mol Mutagen* 2010, **51**:635-645.

52. Fekairi S, Scaglione S, Chahwan C, Taylor ER, Tissier A, Coulon S, Dong MQ, Ruse C, Yates JR 3rd, Russell P *et al.*: **Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases.** *Cell* 2009, **138**:78-89.
53. Munoz IM, Hain K, Declais AC, Gardiner M, Toh GW, Sanchez-Pulido L, Heuckmann JM, Toth R, Macartney T, Eppink B *et al.*: **Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair.** *Mol Cell* 2009, **35**:116-127.
54. Svendsen JM, Smogorzewska A, Sowa ME, O'Connell BC, Gygi SP, Elledge SJ, Harper JW: **Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair.** *Cell* 2009, **138**:63-77.
55. Crossan GP, van der Weyden L, Rosado IV, Langevin F, Gaillard PH, McIntyre RE, Gallagher F, Kettunen MI, Lewis DY, Brindle K *et al.*: **Disruption of mouse Six4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia.** *Nat Genet* 2011, **43**:147-152.
- This study is of particular interest in the field as it described the first FA mouse model that clearly recapitulates the FA clinical phenotype.
56. Kim Y, Spitz GS, Veturi U, Lach FP, Auerbach AD, Smogorzewska A: **Regulation of multiple DNA repair pathways by the Fanconi anemia protein SLX4.** *Blood* 2013, **121**:54-63.
57. Hodskinson MR, Silhan J, Crossan GP, Garaycochea JI, Mukherjee S, Johnson CM, Scharer OD, Patel KJ: **Mouse SLX4 is a tumor suppressor that stimulates the activity of the nuclease XPF-ERCC1 in DNA crosslink repair.** *Mol Cell* 2014, **54**:472-484.
58. Ouyang J, Garner E, Hallet A, Nguyen HD, Rickman KA, Gill G, Smogorzewska A, Zou L: **Noncovalent interactions with SUMO and ubiquitin orchestrate distinct functions of the SLX4 complex in genome maintenance.** *Mol Cell* 2015, **57**:108-122.
59. Guervilly JH, Takedachi A, Naim V, Scaglione S, Chawhan C, Lovera Y, Despras E, Kuraoka I, Kannouche P, Rosselli F *et al.*: **The SLX4 complex is a SUMO E3 ligase that impacts on replication stress outcome and genome stability.** *Mol Cell* 2015, **57**:123-137.
60. Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, Carter KC, Shell BK, Evans E, de Jong MC *et al.*: **Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease.** *Cell* 1996, **86**:811-822.
61. Clauson C, Scharer OD, Niedernhofer L: **Advances in understanding the complex mechanisms of DNA interstrand cross-link repair.** *Cold Spring Harb Perspect Med* 2013, **3**:a012732.
62. Orelli B, McClendon TB, Tsodikov OV, Ellenberger T, Niedernhofer LJ, Scharer OD: **The XPA-binding domain of ERCC1 is required for nucleotide excision repair but not other DNA repair pathways.** *J Biol Chem* 2010, **285**:3705-3712.
63. MacKay C, Declais AC, Lundin C, Agostinho A, Deans AJ, MacArtney TJ, Hofmann K, Gartner A, West SC, Helleday T *et al.*: **Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2.** *Cell* 2010, **142**:65-76.
64. Kratz K, Schopf B, Kaden S, Sendoel A, Eberhard R, Lademann C, Cannavo E, Sartori AA, Hengartner MO, Jiricny J: **Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents.** *Cell* 2010, **142**:77-88.
65. Smogorzewska A, Desetty R, Saito TT, Schlabach M, Lach FP, Sowa ME, Clark AB, Kunkel TA, Harper JW, Colaiacovo MP *et al.*: **A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair.** *Mol Cell* 2010, **39**:36-47.
66. Liu T, Ghosal G, Yuan J, Chen J, Huang J: **FAN1 acts with FANCD2 to promote DNA interstrand cross-link repair.** *Science* 2010, **329**:693-696.
67. Zhao Q, Xue X, Longrich S, Sung P, Xiong Y: **Structural insights into 5' flap DNA unwinding and incision by the human FAN1 dimer.** *Nat Commun* 2014, **5**:5726.
68. Wang R, Persky NS, Yoo B, Ouerfelli O, Smogorzewska A, Elledge SJ, Pavletich NP: **DNA repair. Mechanism of DNA interstrand cross-link processing by repair nuclease FAN1.** *Science* 2014, **346**:1127-1130.
69. Gwon GH, Kim Y, Liu Y, Watson AT, Jo A, Etheridge TJ, Yuan F, Zhang Y, Carr AM, Cho Y: **Crystal structure of a Fanconi anemia-associated nuclease homolog bound to 5' flap DNA: basis of interstrand cross-link repair by FAN1.** *Genes Dev* 2014, **28**:2276-2290.
70. Chaudhury I, Stroik DR, Sobek A: **FANCD2-controlled chromatin access of the Fanconi-associated nuclease FAN1 is crucial for the recovery of stalled replication forks.** *Mol Cell Biol* 2014, **34**:3939-3954.
71. Trujillo JP, Mina LB, Pujol R, Bogliolo M, Andrieux J, Holder M, Schuster B, Schindler D, Surrallés J: **On the role of FAN1 in Fanconi anemia.** *Blood* 2012, **120**:86-89.
72. Zhou W, Otto EA, Cluckey A, Airik R, Hurd TW, Chaki M, Diaz K, Lach FP, Bennett GR, Gee HY *et al.*: **FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair.** *Nat Genet* 2012, **44**:910-915.
73. Segui N, Mina LB, Lazaro C, Sanz-Pamplona R, Pons T, Navarro M, Bellido F, Lopez-Doriga A, Valdes-Mas R, Pineda M *et al.*: **Germline mutations in FAN1 cause hereditary colorectal cancer by impairing DNA repair.** *Gastroenterology* 2015. (Epub ahead of print).
74. Klein Douwel D, Boonen RA, Long DT, Szypowska AA, Raschle M, Walter JC, Knipscheer P: **XPF-ERCC1 acts in unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4.** *Mol Cell* 2014, **54**:460-471.
75. Langevin F, Crossan GP, Rosado IV, Arends MJ, Patel KJ: **Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice.** *Nature* 2011, **475**:53-58.
- In this seminal paper the authors identified, for the first time, naturally generated aldehydes as a possible physiological source of DNA damage linking alcohol metabolism to the FA clinical phenotype.
76. Rosado IV, Langevin F, Crossan GP, Takata M, Patel KJ: **Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway.** *Nat Struct Mol Biol* 2011, **18**:1432-1434.
77. Garaycochea JI, Crossan GP, Langevin F, Daly M, Arends MJ, Patel KJ: **Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function.** *Nature* 2012, **489**:571-575.
- This study shows in adult FA mice that hematopoietic stem and progenitor cells are highly sensitive to aldehydes providing a mechanism for BMF in FA patients.
78. Oberbeck N, Langevin F, King G, de Wind N, Crossan GP, Patel KJ: **Maternal aldehyde elimination during pregnancy preserves the fetal genome.** *Mol Cell* 2014, **55**:807-817.
79. Hira A, Yabe H, Yoshida K, Okuno Y, Shiraishi Y, Chiba K, Tanaka H, Miyano S, Nakamura J, Kojima S *et al.*: **Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients.** *Blood* 2013, **122**:3206-3209.
80. Chen CH, Ferreira JC, Gross ER, Mochly-Rosen D: **Targeting aldehyde dehydrogenase 2: new therapeutic opportunities.** *Physiol Rev* 2014, **94**:1-34.
81. Ruiz S, Fernandez-Capetillo O: **The maternal side of Fanconi anemia.** *Mol Cell* 2014, **55**:803-804.
82. Parmar K, D'Andrea AD: **Stressed out: endogenous aldehydes damage hematopoietic stem cells.** *Cell Stem Cell* 2012, **11**:583-584.
83. Yokoyama A, Yokoyama T, Brooks PJ, Mizukami T, Matsui T, Kimura M, Matsushita S, Higuchi S, Maruyama K: **Macrocytosis, macrocytic anemia, and genetic polymorphisms of alcohol dehydrogenase-1B and aldehyde dehydrogenase-2 in Japanese alcoholic men.** *Alcohol Clin Exp Res* 2014, **38**:1237-1246.
84. Gluckman E, Broxmeyer HA, Auerbach AD, Friedman HS, Douglas GW, Devergie A, Esperou H, Thierry D, Socie G, Lehn P *et al.*: **Hematopoietic reconstitution in a patient with Fanconi's**

- anemia by means of umbilical-cord blood from an HLA-identical sibling. *N Engl J Med* 1989, **321**:1174-1178.
85. Yoder MC: **Cord blood banking and transplantation: advances and controversies.** *Curr Opin Pediatr* 2014, **26**:163-168.
 86. Kapelushnik J, Or R, Slavin S, Nagler A: **A fludarabine-based protocol for bone marrow transplantation in Fanconi's anemia.** *Bone Marrow Transplant* 1997, **20**:1109-1110.
 87. Peffault de Latour R, Porcher R, Dalle JH, Aljurf M, Korthof ET, Svahn J, Willemze R, Barrenetxea C, Mialou V, Soulier J *et al.*: **Allogeneic hematopoietic stem cell transplantation in Fanconi anemia: the European Group for Blood and Marrow Transplantation experience.** *Blood* 2013, **122**:4279-4286.
 88. Verlinsky Y, Rechitsky S, Schoolcraft W, Strom C, Kuliev A: **Preimplantation diagnosis for Fanconi anemia combined with HLA matching.** *JAMA* 2001, **285**:3130-3133.
 89. MacMillan ML, Hughes MR, Agarwal S, Daley GQ: **Cellular therapy for Fanconi anemia: the past, present, and future.** *Biol Blood Marrow Transplant* 2011, **17**:S109-S114.
 90. Kahraman S, Beyazyurek C, Yesilipek MA, Ozturk G, Ertem M, Anak S, Kansoy S, Aksoylar S, Kuskonmaz B, Oniz H *et al.*: **Successful haematopoietic stem cell transplantation in 44 children from healthy siblings conceived after preimplantation HLA matching.** *Reprod Biomed Online* 2014, **29**:340-351.
 91. Trujillo J, Surrallés J: **Savior siblings and Fanconi anemia: analysis of success rates from the family's perspective.** *Genet Med* 2015. (Epub ahead of print).
 92. Tolar J, Becker PS, Clapp DW, Hanenberg H, de Heredia CD, Kiem HP, Navarro S, Qasba P, Rio P, Schmidt M *et al.*: **Gene therapy for Fanconi anemia: one step closer to the clinic.** *Hum Gene Ther* 2012, **23**:141-144.
 93. Liu JM, Kim S, Read EJ, Futaki M, Dokal I, Carter CS, Leitman SF, Pensiero M, Young NS, Walsh CE: **Engraftment of hematopoietic progenitor cells transduced with the Fanconi anemia group C gene (FANCC).** *Hum Gene Ther* 1999, **10**:2337-2346.
 94. Croop JM, Cooper R, Fernandez C, Graves V, Kreissman S, Hanenberg H, Smith FO, Williams DA: **Mobilization and collection of peripheral blood CD34+ cells from patients with Fanconi anemia.** *Blood* 2001, **98**:2917-2921.
 95. Kelly PF, Radtke S, von Kalle C, Balciik B, Bohn K, Mueller R, Schuesler T, Haren M, Reeves L, Cancelas JA *et al.*: **Stem cell collection and gene transfer in Fanconi anemia.** *Mol Ther* 2007, **15**:211-219.
 96. Jacome A, Navarro S, Rio P, Yanez RM, Gonzalez-Murillo A, Lozano ML, Lamana ML, Sevilla J, Olive T, Diaz-Heredia C *et al.*: **Lentiviral-mediated genetic correction of hematopoietic and mesenchymal progenitor cells from Fanconi anemia patients.** *Mol Ther* 2009, **17**:1083-1092.
 97. Gonzalez-Murillo A, Lozano ML, Alvarez L, Jacome A, Almarza E, Navarro S, Segovia JC, Hanenberg H, Guenechea G, Bueren JA *et al.*: **Development of lentiviral vectors with optimized transcriptional activity for the gene therapy of patients with Fanconi anemia.** *Hum Gene Ther* 2010, **21**:623-630.
 98. Rio P, Meza NW, Gonzalez-Murillo A, Navarro S, Alvarez L, Surrallés J, Castella M, Guenechea G, Segovia JC, Hanenberg H *et al.*: **In vivo proliferation advantage of genetically corrected hematopoietic stem cells in a mouse model of Fanconi anemia FA-D1.** *Blood* 2008, **112**:4853-4861.
 99. May M: **Mutations to the rescue.** *Nat Med* 2011, **17**:405-407.
 100. Raya A, Rodriguez-Piza I, Guenechea G, Vassena R, Navarro S, Barrero MJ, Consiglio A, Castella M, Rio P, Sleep E *et al.*: **Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells.** *Nature* 2009, **460**:53-59.
Raya and co-workers showed in this paper that on correction of the genetic defect, somatic cells from FA patients can be reprogrammed to iPS cells to generate disease-free haematopoietic progenitors.
 101. Muller LU, Milsom MD, Harris CE, Vyas R, Brumme KM, Parmar K, Moreau LA, Schambach A, Park IH, London WB *et al.*: **Overcoming reprogramming resistance of Fanconi anemia cells.** *Blood* 2012, **119**:5449-5457.
 102. Yung SK, Tilgner K, Ledran MH, Habibollah S, Neganova I, Singhapol C, Saretzki G, Stojkovic M, Armstrong L, Przyborski S *et al.*: **Brief report: human pluripotent stem cell models of Fanconi anemia deficiency reveal an important role for Fanconi anemia proteins in cellular reprogramming and survival of hematopoietic progenitors.** *Stem Cells* 2013, **31**:1022-1029.
 103. Liu GH, Suzuki K, Li M, Qu J, Montserrat N, Tarantino C, Gu Y, Yi F, Xu X, Zhang W *et al.*: **Modelling Fanconi anemia pathogenesis and therapeutics using integration-free patient-derived iPSCs.** *Nat Commun* 2014, **5**:4330.
Liu and coworker generated integration-free FA IPS cells to model FA pathogenesis and therapeutics.
 104. Rio P, Banos R, Lombardo A, Quintana-Bustamante O, Alvarez L, Garate Z, Genovese P, Almarza E, Valeri A, Diez B *et al.*: **Targeted gene therapy and cell reprogramming in Fanconi anemia.** *EMBO Mol Med* 2014, **6**:835-848.
 105. Navarro S, Moleiro V, Molina-Estevez FJ, Lozano ML, Chinchon R, Almarza E, Quintana-Bustamante O, Mostoslavsky G, Maetzig T, Galla M *et al.*: **Generation of iPSCs from genetically corrected Brca2 hypomorphic cells: implications in cell reprogramming and stem cell therapy.** *Stem Cells* 2014, **32**:436-446.
 106. Marion RM, Strati K, Li H, Murga M, Blanco R, Ortega S, Fernandez-Capetillo O, Serrano M, Blasco MA: **A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity.** *Nature* 2009, **460**:1149-1153.
 107. Chlon TM, Hoskins EE, Mayhew CN, Wikenheiser-Brokamp KA, Davies SM, Mehta P, Myers KC, Wells JM, Wells SI: **High-risk human papillomavirus E6 protein promotes reprogramming of Fanconi anemia patient cells through repression of p53 but does not allow for sustained growth of induced pluripotent stem cells.** *J Virol* 2014, **88**:11315-11326.
 108. Naldini L: **Ex vivo gene transfer and correction for cell-based therapies.** *Nat Rev Genet* 2011, **12**:301-315.
 109. Osborn MJ, Gabriel R, Webber BR, DeFeo AP, McElroy AN, Jarjour J, Starker CG, Wagner JE, Joung JK, Voytas DF *et al.*: **Fanconi anemia gene editing by the CRISPR/Cas9 system.** *Hum Gene Ther* 2015, **26**:114-126.
This reference along with [103*,104] first employed genome editing technologies to genetically correct FA in human FA fibroblasts offering novel therapeutic strategies to treat the disease.
 110. Hartmann L, Neveling K, Borkens S, Schneider H, Freund M, Grassman E, Theiss S, Wawer A, Burdach S, Auerbach AD *et al.*: **Correct mRNA processing at a mutant TT splice donor in FANCC ameliorates the clinical phenotype in patients and is enhanced by delivery of suppressor U1 snRNAs.** *Am J Hum Genet* 2010, **87**:480-493.