



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN CIENCIAS BIOMÉDICAS
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

**IDENTIFICACIÓN DE MUTACIONES EN LOS GENES *FANCL*, EN UN GRUPO DE
PACIENTES MEXICANOS CON ANEMIA DE FANCONI**

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA

PEDRO VICENTE REYES JIMÉNEZ

DIRECTOR DE TESIS

DRA. SARA FRÍAS
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

COMITÉ TUTOR

DRA. MARÍA TERESA VILLARREAL MOLINA
INSTITUTO NACIONAL DE MEDICINA GENÓMICA (INMEGEN)
DR. JORGE MELÉNDEZ ZAJGLA
INSTITUTO NACIONAL DE MEDICINA GENÓMICA (INMEGEN)

MÉXICO, D. F. MAYO DE 2023



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS ACADÉMICOS.

Al Posgrado en Ciencias Biomédicas de la Universidad Nacional Autónoma de México (UNAM), por brindarme la gran oportunidad de crecer académica y profesionalmente.

A la Dra Sara Frías, por permitirme participar en este proyecto y depositar su confianza en mi persona para desarrollarlo. Gracias por su profesionalismo, ética de trabajo, pasión por la ciencia, y amistad.

A mi Comité Tutor, integrado por la Dra. María Teresa Villarreal Molina y el Dr. Jorge Meléndez Zajgla, por su acompañamiento y crítica constructiva durante el transcurso del proyecto. Su orientación oportuna resultó fundamental para alcanzar los objetivos planteados.

Al Instituto Nacional de Pediatría, a través del financiamiento del proyecto de investigación titulado "Caracterización de mutaciones en los genes *FANC*, en un grupo de pacientes mexicanos con Anemia de Fanconi y su correlación con el potencial proliferativo de células troncales hematopoyéticas", registrado bajo el número INP-041-2014.

Al Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCyT), por la asignación de la beca con número de registro 899784, para desarrollar mi Doctorado en el Programa de Ciencias Biomédicas de la UNAM. Asimismo, por la asignación de fondos para el proyecto "Relevancia de la vía FA/BRCA y la reparación del daño al DNA en la apoptosis de las células troncales hematopoyéticas" (SEP-CONACYT 243102), así como también a través del Fondo Sectorial de Investigación en Salud y Seguridad Social (FOSISS), al financiar el proyecto FOSISS-233721. A la Dirección General de Asuntos del Personal Académico de la UNAM, por financiar el proyecto proyecto IN205120 (PAPIIT).

ABREVIATURAS.

AA: anemia aplásica.

ACMG-AMP: Colegio Americano de Genética Médica y Asociación de Patología Molecular (*American College of Medical Genetics and Association for Molecular Pathology*).

AF: anemia de Fanconi.

CMN: células mononucleares.

CNV: variante de número de copias (*Copy Number Variant*).

CTPH: células troncales y progenitoras hematopoyéticas.

DNA: ácido desoxirribonucleico.

DNAg: DNA genómico.

DDR: respuesta al daño al DNA (*DNA damage response*).

DEB: diepoxibutano.

DSB: ruptura de doble hebra (*Double Strand Break*).

FA/BRCA: vía de reparación de Anemia de Fanconi/Cáncer de mama (*Fanconi Anemia/Breast Cancer*).

FCR: fuerza centrífuga relativa.

FM: falla medular.

GATK: conjunto de herramientas para el análisis del genoma (*Genome Analysis Toolkit*).

ICL: enlace covalente intercatenario (*Interstrand crosslink*).

I-D2: complejo proteico FANCI-FANCD2.

INP: Instituto Nacional de Pediatría.

LMA: leucemia mieloide aguda.

LSCH: tramos largos y contiguos de homocigocidad (*Long contiguous stretches of homozygosity*).

MLPA: amplificación de sondas dependientes de ligandos múltiples (*Multiplex ligation-dependent probe amplification*).

MMC: mitomicina C.

MO: médula ósea.

NER: vía de reparación por escisión de nucleótidos (*Nucleotide excision repair*).

NHEJ: reparación por unión de extremos no homólogos (*Non homologous end joining*).

PCR: reacción en cadena de la polimerasa.

PHENOS: acrónimo que engloba anomalías de la Pigmentación, cabeza pequeña (del inglés *small Head*), ojos pequeños (del inglés *small Eyes*), anomalías del Sistema Nervioso Central, alteraciones Otológicas y estatura baja (del inglés *Short stature*).

RH: recombinación homóloga.

RNV: recién nacidos vivos.

SMD: síndromes mielodisplásicos.

SNG: secuenciación de nueva generación.

SP: sangre periférica.

TLS: síntesis translesión (*Translesion synthesis*).

TNF- α : factor de necrosis tumoral alfa.

TPR: repetidos de tetratricopéptidos (*Tetratricopeptide repeats*).

VACTERL-H: acrónimo que engloba alteraciones Vertebrales, Anales, Cardíacas, fístula Traqueo-esofágica, atresia Esofágica, Renales, de extremidades superiores (del inglés *upper Limb*), e Hidrocefalia.

VP: variantes patogénicas.

VUS: variante de significado incierto (*Variant of uncertain significance*).

Nota: Algunas de las abreviaturas se han mantenido en inglés a lo largo del texto, debido a que son ampliamente utilizadas dentro de la comunidad científica, o ante la ausencia de una abreviatura equivalente y estandarizada en idioma español.

Tabla de contenido.

1. RESUMEN.	1
<u>ABSTRACT.</u>	2
2. INTRODUCCIÓN.	3
2.1 LA ANEMIA DE FANCONI. GENERALIDADES.	3
2.2 EPIDEMIOLOGÍA.....	3
2.3 FENOTIPO CLÍNICO EN LA ANEMIA DE FANCONI.	4
2.3.1 Características del desarrollo físico.	4
2.3.2 Características hematológicas	5
2.3.3 Predisposición a cáncer.	6
2.4 FENOTIPO CELULAR.	7
2.5 DIAGNÓSTICO DE LA ANEMIA DE FANCONI.....	8
2.6 ETIOLOGÍA DE LA ANEMIA DE FANCONI.	10
2.6.1 La vía FA/BRCA.	10
2.6.2 Heterogeneidad genética en la Anemia de Fanconi.	13
2.7 CARACTERIZACIÓN MOLECULAR DE VARIANTES PATOGENICAS EN LA ANEMIA DE FANCONI.	16
2.8 APROXIMACIONES A LA CORRELACIÓN GENOTIPO-FENOTIPO EN LA ANEMIA DE FANCONI.	17
3. PLANTEAMIENTO DEL PROBLEMA.	19
4. PREGUNTA DE INVESTIGACIÓN.....	19
5. JUSTIFICACIÓN.....	20
6. HIPÓTESIS.....	20
7. OBJETIVOS.	21
7.1 General	21
7.2 Específicos.	21
8. METODOLOGÍA.....	22
8.1 Diseño del estudio.....	22
8.2 Población de estudio.....	22
8.3 Diseño experimental.	22
8.4 Técnicas empleadas.	25
8.5 Determinación de patogenicidad de variantes candidatas.....	31
8.6 Recopilación de datos clínicos de los pacientes con Anemia de Fanconi.....	32

8.7. Determinación estadística de las correlaciones	32
8.8 Análisis de células troncales y progenitoras hematopoyéticas de pacientes con Anemia de Fanconi a través de citometría de flujo.	33
9. RESULTADOS.	34
9.1 Breve descripción de la población de estudio.....	34
9.2 Resultados que cubren el objetivo particular 1.....	34
9.3 Resultados que cubren el objetivo particular 2.....	38
9.4. Resultados que cubren el objetivo particular 3.....	48
9.5 Resultados que cubren el objetivo particular 4.....	55
10. DISCUSIÓN.....	61
11. CONCLUSIONES.....	71
12. LIMITACIONES Y PERSPECTIVAS.	72
13. REFERENCIAS.	73
14. ANEXOS.....	85

1. RESUMEN.

La Anemia de Fanconi (AF) es una enfermedad hereditaria, clínica y genéticamente heterogénea, que se relaciona con un defecto en la reparación del DNA; está asociada a alteraciones del desarrollo variables, pancitopenia progresiva y predisposición a cáncer. Se han identificado 22 genes causales de la AF (conocidos como genes *FANC*), en los que se han reportado variantes patogénicas (VP) bialélicas y de origen germinal. Estas VP conducen al fenotipo de la AF por herencia autosómica recesiva, con excepción de *FANCB* y *FANCR* (*RAD51*), que se heredan de forma recesiva ligada al X y autosómico dominante, respectivamente. La consecuencia común de estas VP a nivel celular es la inestabilidad cromosómica, característica en la que se basa el diagnóstico citogenético de la enfermedad y que refleja el funcionamiento deficiente de una ruta de reparación del DNA denominada FA/BRCA. La actividad principal de esta vía es la resolución de enlaces covalentes intercatenarios. La gran heterogeneidad genética de la AF dificulta la implementación de técnicas de detección de VP clásicamente utilizadas en otras enfermedades de herencia mendeliana. En México no se ha determinado la frecuencia de estas VP. El objetivo principal de este trabajo consistió en identificar y caracterizar VP en los genes *FANC*, en un grupo de pacientes mexicanos con AF y realizar una correlación genotipo-fenotipo. Para la búsqueda de estas variantes se implementó un diseño experimental en el que se incluyeron la amplificación de sondas múltiple dependiente de ligación (MLPA), secuenciación de nueva generación (SNG) y microarreglos de alta densidad. A través de esta estrategia de genotipificación, se detectaron las variantes responsables de la AF en 41 pacientes, pertenecientes a 33 familias. *FANCA* fue el gen afectado con mayor frecuencia en ellos (56%), seguido por *FANCB* (9.75%), y por *FANCC*, *FANCG* y *FANCF* (7.3% cada uno). Se encontraron siete tipos de variantes; las más comunes fueron las de sentido erróneo y las de cambio de marco de lectura, que fueron identificadas en la misma proporción (21.95%). La variante con más alelos afectados (siete) en este estudio fue *FANCA*: c.2535_2536delCT. Asimismo, se determinaron dos variantes con efecto fundador en nuestra población de pacientes: *FANCC*: c.67delG en una comunidad Menonita (Tamaulipas) y *FANCG*: c.511-3_511-2delCA en una comunidad Mixe (Oaxaca), siendo esta última la primera variante de este tipo de ancestría mexicana. Tres pacientes (7.31%) desarrollaron procesos neoplásicos. Las alteraciones físicas, descritas a través de los acrónimos VACTERL-H y PHENOS, así como las hematológicas (evaluadas a través de la clasificación de su falla medular) fueron utilizadas para describir el fenotipo de los pacientes. En tal sentido, las anomalías físicas más recurrentes fueron talla baja (44.4%), defectos en las extremidades (56%) y cambios en la pigmentación de la piel (65.85%); estos últimos asociados a variantes nulas. Respecto a la severidad de la falla medular, su ausencia o presentación en forma leve o moderada, prevalece en los pacientes con genotipo *FANCA*, cuando se contrasta con pacientes con otros genotipos. Para profundizar en las implicaciones de una vía FA/BRCA deficiente sobre las células troncales y progenitoras hematopoyéticas (CTPH, identificadas a través de la presencia del marcador CD34), se utilizaron ensayos de citometría de flujo para estimar su abundancia y grado de daño genómico, a partir de muestras de médula ósea (MO) y sangre periférica (SP) de 12 pacientes. Los resultados señalan que los afectados con AF tienen un número reducido de CTPH en la MO, así como una tendencia de estas células a permanecer en SP, cuando fueron comparadas con muestras de individuos sanos. La fracción de CTPH presente en SP presentó mayor acumulación de daño genómico. La identificación oportuna y sistemática de las variantes responsables del fenotipo en la AF en el país es un prerrequisito para el asesoramiento genético de las familias, manejo clínico de los pacientes (en especial para el trasplante de médula ósea), el descubrimiento de asociaciones genotipo-fenotipo y para conocer la composición genómica de la AF en la población mexicana.

1. ABSTRACT.

Fanconi anemia (FA) is a clinically and genetically heterogeneous inherited disease related to a defect in DNA repair; it is associated with various developmental disorders, progressive pancytopenia, and a predisposition to cancer. Twenty-two FA-causing genes (known as *FANC* genes) have been identified, in which bi-allelic germline pathogenic variants (PV) have been reported. These PV lead to the FA phenotype by autosomal recessive inheritance, with the exception of *FANCB* and *FANCR* (*RAD51*), which are inherited in an X-linked recessive and autosomal dominant fashion, respectively. The main consequence of these PV at the cellular level is chromosomal instability, a feature that underlies the cytogenetic diagnosis of the disease and reflects the dysfunction of a DNA repair pathway called FA/BRCA. The key activity of this pathway is the resolution of interstrand cross-links. The great genetic heterogeneity of FA hampers the use of PV detection approaches classically used in other inherited diseases. The frequency of these PV has not been determined in Mexico. The main objective of this work was to identify and characterize PV in the *FANC* genes in a group of Mexican patients with FA and to perform genotype-phenotype correlation. For this task, we followed an experimental design that included multiplex ligation-dependent probe amplification (MLPA), next-generation sequencing (NGS) and high-density microarrays. Through this genotyping strategy, genetic variants responsible for FA were detected in 41 patients, belonging to 33 families. *FANCA* was the most frequently affected gene (56%) in them, followed by *FANCB* (9.75%), and by *FANCC*, *FANCG* and *FANCF* (7.3% each). Seven types of variants were found; the most common were missense and frameshift variants, which were identified in the same proportion (21.95%). The variant with the most affected alleles (seven) in this study was *FANCA*: c.2535_2536delCT. Two variants with founder effect were also identified in our patient group: *FANCC*: c.67delG in a Mennonite community (Tamaulipas) and *FANCG*: c.511-3_511-2delCA in a Mixe community (Oaxaca), which represents the first variant of this type of Mexican ancestry. Three of these patients (7.31%) developed neoplastic processes. Physical alterations, described through VACTERL-H and PHENOS acronyms, as well as hematological alterations (assessed through the classification of their bone marrow failure) were used to describe the phenotype of the patients. The most recurrent physical abnormalities were short stature (44.4%), limb defects (56%), and skin pigmentation changes (65.85%), the latter associated with the presence of null variants. Regarding the severity of bone marrow failure, its absence or mild to moderate types are prevalent in patients with *FANCA* genotype, when contrasted with patients with other genotypes. To further investigate the scope of a deficient FA/BRCA pathway on hematopoietic stem and progenitor cells (HSPCs, identified by the presence of the CD34 surface marker), flow cytometry assays were used to estimate their abundance and degree of genomic damage in bone marrow (BM) and peripheral blood (PB) samples from 12 patients. The results indicate that those affected with FA have a reduced number of HSPCs in BM, in addition to a tendency for these cells to remain in SP when compared to samples from healthy individuals. The HSCs fraction located in SP showed a higher accumulation of genomic damage. The prompt and systematic identification of the variants responsible for the FA phenotype in the country is a prerequisite for genetic counseling of families, clinical management of patients (especially for bone marrow transplantation), the discovery of genotype-phenotype associations, and understanding the genomic composition of FA in the Mexican population.

2. INTRODUCCIÓN.

2.1 LA ANEMIA DE FANCONI. GENERALIDADES.

La Anemia de Fanconi (AF) es una enfermedad hereditaria, clínica y genéticamente heterogénea, que se relaciona con un defecto en la reparación del DNA asociado a alteraciones del desarrollo variables, pancitopenia progresiva y predisposición a cáncer (Nalepa y Clapp, 2018).

La AF fue descrita por primera vez en 1927 por el pediatra suizo Guido Fanconi, en una familia de cinco hermanos, tres de los cuales fallecieron debido a sangrados espontáneos y anemia severa, asociadas a microcefalia congénita, alteraciones en la piel, hemorragias cutáneas e hipoplasia testicular (Fanconi, 1927). Estas observaciones iniciales formaron parte de los criterios aplicados para el diagnóstico de la enfermedad por muchos años, hasta que se describió que la AF estaba relacionada a rupturas cromosómicas espontáneas (Schroeder, 1966). Tres años después, se demostró que las células de estos pacientes eran altamente sensibles ante la exposición de agentes alquilantes, incrementando sus aberraciones cromosómicas luego del tratamiento con estos compuestos (Schuler et al, 1969). Esta característica se utilizó para la estandarización de la prueba diagnóstica de la AF (Esmer et al, 2004, Auerbach, 2009).

Con el desarrollo de la biología molecular, y luego de varias décadas de investigación, en 1992 se descubrió el primer gen causante de la AF, *FANCC* (Strathdee et al, 1992). Posteriormente se describieron a *FANCA* (Fanconi anaemia/Breast cancer consortium, 1996), *FANCG* (de Winter et al, 1998), *FANCE* (de Winter et al, 2000), *FANCF* (de Winter et al, 2000) y *FANCD2* (Timmers et al, 2001). Hasta la fecha, se han identificado 22 genes causales de la AF. En todos ellos se han reportado variantes patogénicas (VP) bialélicas y de origen germinal, que conducen al fenotipo de la AF por herencia autosómica recesiva, con la excepción de *FANCB* (herencia recesiva ligada al X) y *FANCR* (*RAD51*) (herencia autosómico dominante), en las cuales se encuentra afectado un sólo alelo (García de Teresa et al, 2020). La consecuencia común de estas VP es la inestabilidad cromosómica, reflejo del funcionamiento deficiente de una ruta de reparación en el DNA, denominada FA/BRCA (del inglés *Fanconi Anemia/Breast Cancer*) (Merford et al, 2022).

2.2 EPIDEMIOLOGÍA.

La AF tiene una distribución global y ha sido descrita en todos los grupos raciales. En promedio, se calcula que la incidencia de la AF es de 1 por cada 360000 recién nacidos vivos (RNV) (Mamrak et al, 2017). Algunos reportes de registros europeos señalan que la incidencia es de 4-7 por millón de RNV (Gulbis et al, 2010). En Estados Unidos se estima que afecta a 1 por cada 100000 RNV, con una frecuencia de portadores de 1:181; mientras que en Israel es de 1:93 (calculados en el lapso 1989-2000). Sin embargo, varios autores sostienen que la AF está sub-diagnosticada, por lo que se infiere que la frecuencia de portadores debe ser mayor (Rosemberg et al, 2011). La proporción de afectados entre hombres y mujeres es 1.2:1 (Meetei et al, 2004). Hasta principios del siglo XXI, se habían reportado más de 1300 casos en

la literatura, particularmente en pacientes con alteraciones físicas que desarrollaron anemia aplásica o leucemia (Auerbach, 2009).

En México se desconocen los datos epidemiológicos de la AF, sin embargo, en el Laboratorio de Citogenética del Instituto Nacional de Pediatría (INP), considerado como centro de referencia nacional para el diagnóstico de la enfermedad, se diagnostican aproximadamente 10 casos nuevos al año por cada 80 muestras enviadas con sospecha, provenientes de múltiples instituciones de salud de toda la República (García de Teresa et al, 2016).

A pesar de ser catalogada como una enfermedad rara, la AF presenta mayor prevalencia en ciertos grupos poblacionales, debido a VP con efecto fundador, que incrementan la frecuencia de portadores (ver sección 2.6.2). Entre estos grupos podemos mencionar a los judíos Ashkenazi (*FANCC* y *FANCD1/BRCA2*), Afrikáneres (*FANCA*), africanos del Sub-Sahara (*FANCG*), gitanos españoles (*FANCA*), entre otros (Kutler y Auerbach, 2004; Whitney et al, 1993; Tipping et al, 2001; Morgan et al, 2005; Callén et al, 2005).

2.3 FENOTIPO CLÍNICO EN LA ANEMIA DE FANCONI.

2.3.1 Características del desarrollo físico.

El fenotipo físico de los pacientes con AF es muy heterogéneo y puede afectar múltiples sistemas. Dentro de los hallazgos más frecuentes, considerados como “clásicos”, se encuentran talla baja, anomalías del eje radial (que pueden incluir alteraciones o ausencia de pulgares), cambios en la pigmentación de la piel, manchas *café au lait*, facies característica y malformaciones urogenitales (Savage y Walsh, 2018). Las alteraciones en el desarrollo encontradas comúnmente en la enfermedad incluyen aquellas descritas en la asociación **VACTERL-H** (Vertebrales, Anales, Cardíacas, fístula Traqueo-esofágica, atresia Esofágica, Renales, de extremidades (del inglés *Limb*), e Hidrocefalia). Se estima que 5%-30% de los pacientes con AF reúnen al menos 3 de las 8 características necesarias para presentar esta asociación (Alter y Giri, 2016).

Aunque los pacientes con AF exhiban características propias de la asociación VACTERL-H, no es específica de la AF. Debido a ello, en el año 2016 se propuso el acrónimo **PHENOS** (anomalías de la Pigmentación, cabeza pequeña (del inglés *small Head*), ojos pequeños (del inglés *small Eyes*), anomalías del Sistema Nervioso Central, alteraciones Otológicas y estatura baja (del inglés *Short stature*) para complementar la identificación de pacientes, ya que estos hallazgos también son frecuentes en la enfermedad y no forman parte de la asociación VACTERL-H (Alter y Giri, 2016; Fiesco-Roa et al, 2019).

Por otra parte, las anomalías endocrinas ocurren en un alto porcentaje de los pacientes que presentan alteraciones en el desarrollo; estas incluyen hiperinsulinemia, insuficiencia de la hormona del crecimiento e hipotiroidismo (Wajnrajch et al, 2001). Los aspectos reproductivos también se encuentran comprometidos, debido a la falla en la espermatogénesis en los hombres, frecuentemente acompañada de azoospermia no obstructiva y a la fertilidad reducida en las mujeres, que puede manifestarse como insuficiencia ovárica primaria (Tsui y Crimani, 2019).

A pesar del gran número de manifestaciones físicas mencionadas, entre 20%-25% de los pacientes con AF presenta pocas o ninguna de estas características, por lo que la ausencia de anomalías no descarta el diagnóstico de la enfermedad (Kutler et al, 2003, Fiesco-Roa et al, 2019). Por ello, la presunción diagnóstica de la AF, apoyada únicamente en hallazgos físicos y de gabinete, no excluye la presencia de otros síndromes, dada la diversidad y similitud de sus manifestaciones. Entre estas enfermedades, consideradas como diagnósticos diferenciales de la AF, podemos mencionar el Síndrome de Seckel, Síndrome de Nijmegen, Síndrome de Holt-Oram, Síndrome de trombocitopenia con ausencia de radio, Anemia de Diamond-Blackfan, Disqueratosis congénita, entre otros (Neveling et al, 2009; Auerbach, 2009).

2.3.2 Características hematológicas.

Los pacientes con AF generalmente desarrollan algún grado de disfunción de la médula ósea (MO). La falla medular (FM), que se refiere a la producción disminuida de uno o más de los linajes hematopoyéticos principales, es la característica más encontrada en los pacientes. La edad promedio de aparición de la FM es a los 7 años, aunque el rango de inicio es variable (Rosemberg et al, 2011, Alter et al, 2018). En la mayoría de los casos, la falla comienza con trombocitopenia, seguida de leucopenia. El desarrollo progresivo de la FM culmina en pancitopenia. Cuando la FM se agrava, avanza hasta la aparición de la anemia aplásica (AA), condición en la que la pancitopenia periférica se asocia a una MO hipocelular (<25%).

La severidad de la FM puede clasificarse en tres grandes grupos, en base al grado de citopenia presentada, siendo estas persistentes y no secundarias a otra causa tratable (Sroka et al, 2020; ver Tabla 1).

Tabla 1. Severidad de la falla medular

	Leve	Moderada (hipoplásica o anemia aplásica)	Severa (anemia aplásica severa)
Conteo absoluto de neutrófilos	<1,500/mm ³	<1,000/mm ³	<500/mm ³
Conteo de plaquetas	150,000-50,000/mm ³	<50,000/mm ³	<30,000/mm ³
Nivel de hemoglobina (Hb)	≥8 g/dL*	<8 g/dL	<8 g/dL

En la Tabla: dL: decilitro, g: gramos, mm³: milímetro cúbico, * menor de lo normal para la edad pero >8 g/dL (tomada de Sroka et al, 2020).

Se estima que tres de cada cuatro pacientes con AF desarrollará algún grado de FM durante la primera década de vida (Rosenberg et al, 2003). Aunque muchos de los afectados prosiguen hacia la AA, otros mantienen niveles intermedios en sus conteos sanguíneos por años, e incluso décadas. El diagnóstico y seguimiento de la AA, en caso de presentarse, debe incluir conteo sanguíneo completo, conteo de reticulocitos y una biopsia de MO. Estos parámetros frecuentemente se integran para clasificar la gravedad de la AA, de acuerdo a

los criterios propuestos por Camitta (Camitta et al, 1982, DeZern y Brodsky, 2011). Asimismo, la estimación del porcentaje de CTPH en la MO, que puede evaluarse a través de la identificación del antígeno de superficie CD34 (presente en las células sanguíneas más primitivas), también resulta de gran utilidad para esta clasificación (Matsui et al, 2006).

La AF se considera como el síndrome de FM hereditaria más frecuente, y la deficiencia eventual de todos los linajes sanguíneos en los pacientes sugiere un deterioro en la funcionalidad de las células troncales y progenitoras hematopoyéticas (CTPH). Cabe destacar que existen reportes que indican que los pacientes con AF con alteraciones severas en el desarrollo también manifiestan frecuentemente la FM. Esta observación, sumada a evidencias que soportan que el defecto en las CTPH comienza en el útero, propone que las anomalías en el desarrollo y la depleción de las CTPH pudiesen ser consecuencia de la respuesta inadecuada ante el daño en el DNA durante las primeras etapas embrionarias (Garaycochea y Patel, 2013).

Diversos estudios sugieren que el agotamiento de las CTPH durante la FM en la AF se produce debido a una MO proinflamatoria, mediada por la liberación de citocinas como INF- γ y TNF- α , y a la toxicidad inducida por aldehídos endógenos. Estos factores, aunados a la deficiencia en la vía FA-BRCA, promueven la hiperactivación de vías de reparación alternativas propensas a error, en respuesta al daño en el DNA (Ceccaldi et al, 2016, Savage y Walsh, 2018). La convergencia de tales eventos conduce a la muerte de las CTPH por apoptosis, mediada por la hiperactivación de las vías de p53/p21 y TGF- β (Garaycochea et al, 2018, Walter et al, 2015). Por otra parte, las señales de supresión de crecimiento son contrarrestadas a su vez por la sobreexpresión de MYC, que fomenta la proliferación de una fracción de las CTPH a expensas del daño en el DNA, permitiendo que los pacientes sobrevivan (Rodríguez et al, 2020, Rodríguez et al, 2022).

Estas evidencias respaldan la hipótesis que plantea que la inestabilidad genómica inherente de la AF, en combinación con inflamación crónica, representan una fuerte presión selectiva para la evolución de clonas malignas en la MO de los pacientes y el posible desarrollo de síndromes mielodisplásicos (SMD) y/o leucemia mieloide aguda (LMA) (Peffault de Latour & Soulier, 2016; Cioc et al, 2010). Para monitorear esta progresión, se recomienda la valoración de anomalías morfológicas y citogenéticas en biopsias y aspirados de MO, iniciando al momento del diagnóstico y continuando durante evaluaciones anuales subsecuentes (Sroka et al, 2020).

En los últimos 30 años, la expectativa de vida de los pacientes con AF ha mejorado significativamente cuando se realiza un trasplante de CTPH en la fase aplásica, con donadores relacionados y no relacionados, siendo los desafíos a largo plazo la reducción de endocrinopatías y procesos neoplásicos futuros (Bonfim, 2020).

2.3.3 Predisposición a cáncer.

Se considera que la AF es un síndrome de predisposición a cáncer. Las neoplasias que se presentan en la enfermedad incluyen principalmente SMD y LMA, así como algunos tumores sólidos. La aparición de estas manifestaciones varía de acuerdo a la edad. Referente a la malignización hematológica, el riesgo relativo de desarrollar LMA y SMD en la AF es de 700 y 6000 veces más, respectivamente, comparado con la población general. Se ha estimado que la edad promedio de los pacientes con AF que presentan LMA es de 14 años. Otras

formas de leucemia, como la leucemia linfoblástica aguda, son mucho más raras (Alter, 2014; Dufour, 2017).

Numerosos reportes derivados de análisis citogenéticos en MO han revelado que cambios clonales en los cromosomas uno, tres y siete son los más comunes en la evolución hacia SMD y LMA en pacientes con AF (Alter et al, 2018). Las opciones terapéuticas en estos casos difieren de las empleadas de manera convencional en población sin AF, debido a la hipersensibilidad de las células de los pacientes a los compuestos citotóxicos aplicados en el tratamiento de estas condiciones, recomendándose quimioterapia de intensidad reducida, seguida de trasplante de CTPH (Alter, 2014).

Respecto a los tumores sólidos, el riesgo de desarrollar carcinoma de células escamosas de cabeza y cuello (sobre todo en la cavidad oral) es 500 veces mayor en adolescentes y adultos jóvenes con AF (Kutler et al, 2003). En un estudio reciente, llevado a cabo en el Instituto Nacional de Cáncer de Estados Unidos, la edad promedio de los pacientes con este carcinoma fue de aproximadamente 34 años (Altintas et al, 2023). Este tumor sólido es el más frecuente en la AF, tanto en pacientes trasplantados como en aquellos que no, seguidos por los casos de cáncer de hígado y cerebro (Alter, 2014).

Finalmente, se ha reportado que la presencia de VP en estado heterocigoto en los genes *FANC* puede contribuir al desarrollo de cáncer en individuos sin AF. Por ejemplo, VP en *FANCD1 (BRCA2)* incrementan el riesgo de cáncer de mama hereditario en casi 50%, mientras que el riesgo para cáncer de ovario es de 15%. Porcentajes de predisposición similar para estas neoplasias también se han reportado para variantes en *FANCS (BRCA1)* (70% y 65%, respectivamente); y también, aunque con una frecuencia de VP mucho menor, en *FANCI (BRIP1)* y *FANCN (PALB2)* (Nalepa y Clapp, 2018, Fu et al, 2022).

2.4 FENOTIPO CELULAR.

Aunque se ha observado una amplia variación en la presentación clínica de la AF, el fenotipo celular es altamente constante en los pacientes, sin importar el gen afectado o el tipo de VP. Los principales rasgos que definen a las células con AF incluyen un incremento en las aberraciones cromosómicas espontáneas e hipersensibilidad a los agentes inductores de enlaces cruzados intercatenarios en el DNA (abreviados como ICLs, del inglés *interstrand cross-links*), como la mitomicina C (MMC), el diepoxibutano (DEB) y el cisplatino. Sin embargo, existen otras particularidades que también contribuyen a caracterizarlas, como el arresto del ciclo celular en G2/M, la producción de citocinas proapoptóticas y el acortamiento de sus telómeros, que ha sido vinculado a estrés oxidativo (Helbing-Leclerc et al, 2021).

Usualmente, durante el cultivo de linfocitos derivados de pacientes con sospecha de AF, los linfocitos se tratan con fitohematoglutina, un mitógeno que estimula la división celular. Sin embargo, las células de los pacientes con AF alargan la duración de la fase G2/M del ciclo celular, debido al alto nivel de daño no reparado. Este efecto es aún más pronunciado cuando se tratan con agentes inductores de ICLs, y puede ser utilizado de manera complementaria como indicador adicional en el diagnóstico, utilizando para ello citometría de flujo (Peake y Noguchi, 2022).

Asimismo, se ha demostrado que estas células producen niveles elevados de citocinas pro-apoptóticas. Se ha encontrado sobreproducción del factor de necrosis tumoral alfa (TNF- α) en suero, plasma y a nivel intracelular en pacientes con AF. Esta hiperestimulación,

sumada a la generación alterada de otros factores de crecimiento y citocinas (que incluyen la expresión reducida de interleucina 6 y del factor estimulante de colonias de granulocitos y macrófagos) puede modificar el microambiente hematopoyético y conducir a una exposición sostenida a inhibidores mitogénicos (Du et al, 2013). Estas fuentes de estrés promueven la apoptosis excesiva de las CTPH, en las que se han documentado niveles incrementados del receptor de muerte Fas (CD95), NF-kappa B y de p53. Este último representa un factor crítico en la regulación del ciclo y la muerte celular luego del daño inducido en el DNA, en respuesta al estrés fisiológico, que al estar alterado fomenta la depleción de las CTPH (Cecaldi et al, 2012, Li et al, 2018).

En los pacientes con AF también se ha descrito que los telómeros son más cortos, comparados con personas sanas de la misma edad. Se ha especulado que este acortamiento se asocia al daño oxidativo, y algunos autores proponen que la sobreproducción de especies reactivas en oxígeno interfiere con el potencial reductor celular. El desbalance en la generación de ATP y la acumulación de proteínas oxidantes, que han sido caracterizadas en distintos subgrupos de AF, evidencian una fuerte disfunción mitocondrial, que pudiese representar un mecanismo complementario a los ICLs para la formación de rupturas cromosómicas (Callen et al, 2002, Kumari et al, 2012; Pagano et al, 2021).

Por otra parte, el estrés oxidativo se ha vinculado a la etiología de enfermedades relacionadas a un fenotipo celular de envejecimiento, entre las que se ha incluido la AF. Varios estudios en células deficientes en la vía FA/BRCA han demostrado que presentan características propias de la senescencia, como proliferación reducida, activación del eje p53-p21, p16 y ATM, así como expresión de la β -galactosidasa (Helbling-Le Clerc et al, 2021).

2.5 DIAGNÓSTICO DE LA ANEMIA DE FANCONI.

La AF es una enfermedad que presenta heterogeneidad fenotípica y expresividad variable, que se puede sospechar -pero no diagnosticar- basándose únicamente en las manifestaciones clínicas. Luego de conocerse la gran sensibilidad de las células de los pacientes al efecto inductor de rupturas cromosómicas por parte de agentes productores de ICLs, como la MMC o el DEB, esta característica se ha utilizado rutinariamente para diagnosticar la enfermedad, a través de la “prueba de aberraciones cromosómicas”. Las indicaciones para esta prueba típicamente incluyen a pacientes con anomalías congénitas, con/sin trombocitopenia y/o falla medular (Oostra et al, 2012). El diagnóstico correcto y oportuno a través de esta prueba, considerada como el “estándar de oro”, es fundamental, ya que permite un asesoramiento genético adecuado y una mejor planificación familiar (Merfort et al, 2022).

Durante el ensayo citogenético, los linfocitos T de sangre periférica de un probando se cultivan con MMC o DEB, y posteriormente se cuantifican las aberraciones en los cromosomas metafásicos teñidos con colorante de Giemsa (Oostra et al, 2012). Actualmente, el uso del DEB es más recomendado, ya que genera menor citotoxicidad y su tasa de falsos positivos y falsos negativos es menor (Auerbach, 2015). De forma paralela, también se deben establecer valores de referencia para rupturas cromosómicas en células no tratadas, ya que estos valores presentan una gran variabilidad entre los pacientes. En adición, se cultiva un control positivo (usualmente una línea linfoblastoide con fenotipo AF) y un control negativo (linfocitos de un individuo sano). Típicamente, las metafases examinadas en los cultivos de

un individuo con AF muestran múltiples rupturas y rearrreglos cromatídicos y cromosómicos por célula, que pueden incluir fragmentos acéntricos y figuras radiales (Molina et al, 2022) (Ver Figura 1).

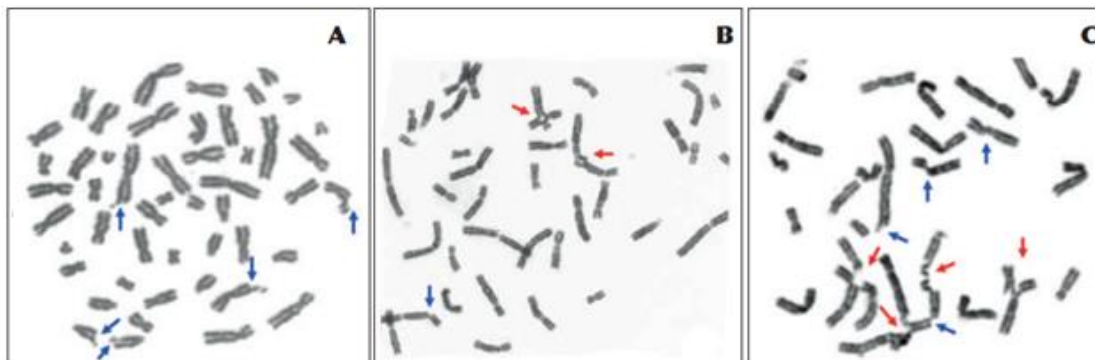


Figura 1. Diagnóstico citogenético de un paciente positivo para Anemia de Fanconi. A. Metafases de un paciente con AF sin tratamiento. Las flechas azules señalan rupturas cromatídicas e isocromatídicas. La frecuencia de estas aberraciones (0.30 ab/cel) es superior a las que se pueden encontrar en un individuo sano. B. Metafase de un paciente positivo para la AF, tratadas con un agente inductor de ICLs (DEB). Las flechas rojas señalan la presencia de figuras de intercambio radial, que son características de la enfermedad. La frecuencia en este caso es de 4.24 ab/cel. C. Metafase de una línea linfoblastoide, utilizada como control positivo para el diagnóstico de AF, en la que se observa un elevado número de aberraciones. En este caso, la frecuencia de aberraciones por célula fue de 10.2 (Tomado y modificado de Molina et al, 2022).

Se estima que el nivel de incremento de estas alteraciones es 10 veces mayor en los cultivos tratados con DEB, respecto a su cultivo sin tratamiento, siendo indicativos de AF (García de Teresa et al, 2020). Sin embargo, en algunos pacientes la prueba puede ser inconclusa, debido a que una subpoblación de las células evaluadas resulta insensible a los agentes clastogénicos, en cuyo caso se sospecha de mosaicismo somático. Esta condición surge debido a diversos eventos (reversión de la variante patogénica, conversión génica, variante de ganancia de función, deleciones/inserciones compensatorias, entre otras), que generan que uno de los dos alelos con la variante causal de la AF pueda producir una proteína funcional a nivel de células troncales o progenitoras linfoides (Dufour y Pierri, 2022). En estos casos, y si la evidencia clínica de AF es consistente, se debe realizar la misma prueba en fibroblastos, obtenidos de una muestra de piel, ya que en este tipo de tejido generalmente no se presenta la reversión de la VP (Sroka et al, 2020).

2.6 ETIOLOGÍA DE LA ANEMIA DE FANCONI.

2.6.1 La vía FA/BRCA.

La vía FA/BRCA es una red de proteínas que mantiene la integridad del genoma, y se activa para proteger a las horquillas de replicación durante la fase de síntesis en el ciclo celular y promover la reparación del DNA (Niraj et al, 2019; García de Teresa et al, 2020). Esta ruta forma parte de los mecanismos de respuesta al daño en las células (abreviado como DDR, del inglés *DNA damage response*) y está compuesta por al menos 22 proteínas, codificadas por los genes *FANC*, además de otras proteínas accesorias. Algunos de estos genes también codifican proteínas de susceptibilidad al cáncer de mama, razón por la cual la vía se conoce como FA/BRCA (Che et al, 2018).

La principal función de la vía FA-BRCA es orquestar la reparación de los ICLs. Estas lesiones son uniones covalentes entre las dos hebras del DNA, que distorsionan la estructura de la doble hélice, impidiendo la replicación y la transcripción al evitar la separación de las dos hebras (Nalepa y Clapp, 2018). La fuente de los ICLs puede ser endógena o exógena. En el primer caso, los ICLs son producidos por aldehídos, especies reactivas de oxígeno y ácido nitroso, entre otras sustancias generadas por el metabolismo celular, capaces de reaccionar con la molécula de DNA. Por otra parte, la fuente exógena está representada por la exposición a diversos agentes ambientales, así como a quimioterapéuticos (en especial aquellos que poseen platino) y ciertas mostazas nitrogenadas (Deans y West, 2011; Garaycochea et al, 2018).

Clásicamente, la vía FA-BRCA suele ser descrita en tres complejos, de acuerdo a su función en la reparación de los ICLs. Las proteínas “*río arriba*”, interactúan para formar el complejo central en respuesta al daño en el DNA, que cataliza la monoubiquitinación del segundo complejo, formado por FANCI y FAND2. Este cambio, a su vez activa a las proteínas “*río abajo*”, que cooperan para la reparación del DNA (Fiesco-Roa et al, 2019).

La reparación de los ICLs ocurre predominantemente durante la fase S del ciclo celular, a través de 4 etapas, que se describen brevemente a continuación (Figura 2):

- a. *Reconocimiento de la lesión*: FANCM, en interacción con FAAP24, MHF1 y MHF2, forma un complejo “*ancla*” que detecta la lesión en el DNA, cuando dos horquillas de replicación convergen en las cercanías de un ICL (Zhang et al, 2015). Este evento asegura el reclutamiento del complejo central (denominado en inglés como “*core*”) hacia los ICLs en la cromatina.
- b. *Reclutamiento del complejo central*: El complejo central es un ensamblaje formado por al menos 8 proteínas (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL y FANCT), además de otras proteínas accesorias (FAAP20 y FAAP100), cuya función principal es la monoubiquitinación del heterodímero compuesto por FANCI y FANCD2. Para catalizar esta reacción, FANCT y FANCL actúan como enzimas E2 ubiquitina conjugante y E3 ubiquitina ligasa, respectivamente (Nalepa y Clapp, 2018, Shakeel et al, 2019). Se ha propuesto que la conformación del complejo central esta mediada por la interacción de 3 módulos o “*subcomplejos*”, cuyo nombre hace alusión a las proteínas que los conforman: BL100, también llamado catalítico, que guía el ensamblaje total del complejo central; CEF, que funciona como sustrato de reconocimiento para la interacción

entre el complejo central y su blanco FANCI-FANCD2; y AG20, que desempeña un rol de conexión estructural o puente entre los otros 2 módulos, además de mediar la localización del complejo central hacia el núcleo (Rodríguez y D'Andrea, 2017; Swiec et al, 2017; van Twest et al, 2017).

- c. *Monoubiquitinación del complejo FANCI-FANCD2*: El heterodímero formado por FANCI y FANCD2 se recluta a los sitios donde se localizan las horquillas estancadas, donde FANCI se trifosforila por la cinasa ATR, estimulando la monoubiquitinación de ambas proteínas por el complejo central. Este evento protege a las horquillas de la degradación nucleolítica y permite el reclutamiento de las proteínas de reparación “rio abajo” (García de Teresa et al, 2020; Tan et al, 2020).
- d. *Remoción del ICL y reparación de lesiones*. El complejo FANCD2-FANCI monoubiquitinado recluta a FANCP/SLX4, una proteína de andamiaje que coordina a las endonucleasas MUS81, SLX1 y FANCO/XPF para realizar el corte del DNA a ambos lados del ICL, desenganchándolo. Como consecuencia, se generan dos intermediarios de reparación en el DNA: un segmento de hebra sencilla, acompañado de un aducto de DNA, que serán reparados a través de síntesis translesión (abreviado como TLS, en la que participa FANCV/REV7) y reparación por escisión de nucleótidos (abreviado como NER, por sus siglas en inglés); y una ruptura de doble hebra, que será reparada por recombinación homóloga, un proceso libre de errores en el que participan FANCD1/BRCA2, FANCN/PALB2, FANCS/BRCA1, FANCI/BRIP1, FANCO/RAD51C, FANCR/RAD51, FANCU/XRCC2 y FANCW/REV3 (García de Teresa et al, 2020; Fiesco-Roa et al, 2022).

Una vez que se cumplen estos pasos, se restaura la actividad de las horquillas y el dímero FANCD2-FANCI es deubiquitinizado por el complejo USP1-UAF1, liberándolo de la cromatina y reestableciendo la síntesis de DNA (van Twest et al, 2017).

La perturbación en la vía FA/BRCA conduce a defectos en la recombinación homóloga, por lo que las células canalizan la reparación de las rupturas de doble hebra preferencialmente hacia vías alternativas, como la unión de extremos no homólogos (abreviada como NHEJ, del inglés *non-homologous end joining*) y potencialmente, la unión de extremos mediada por micro-homología. Ambas vías son propensas a errores y tienen un impacto negativo sobre la integridad genómica, al propiciar la unión errónea entre cromátidas no-hermanas (Juárez-Figueroa et al, 2018; García de Teresa et al, 2020; Fiesco-Roa et al, 2022; Adamo et al, 2010).

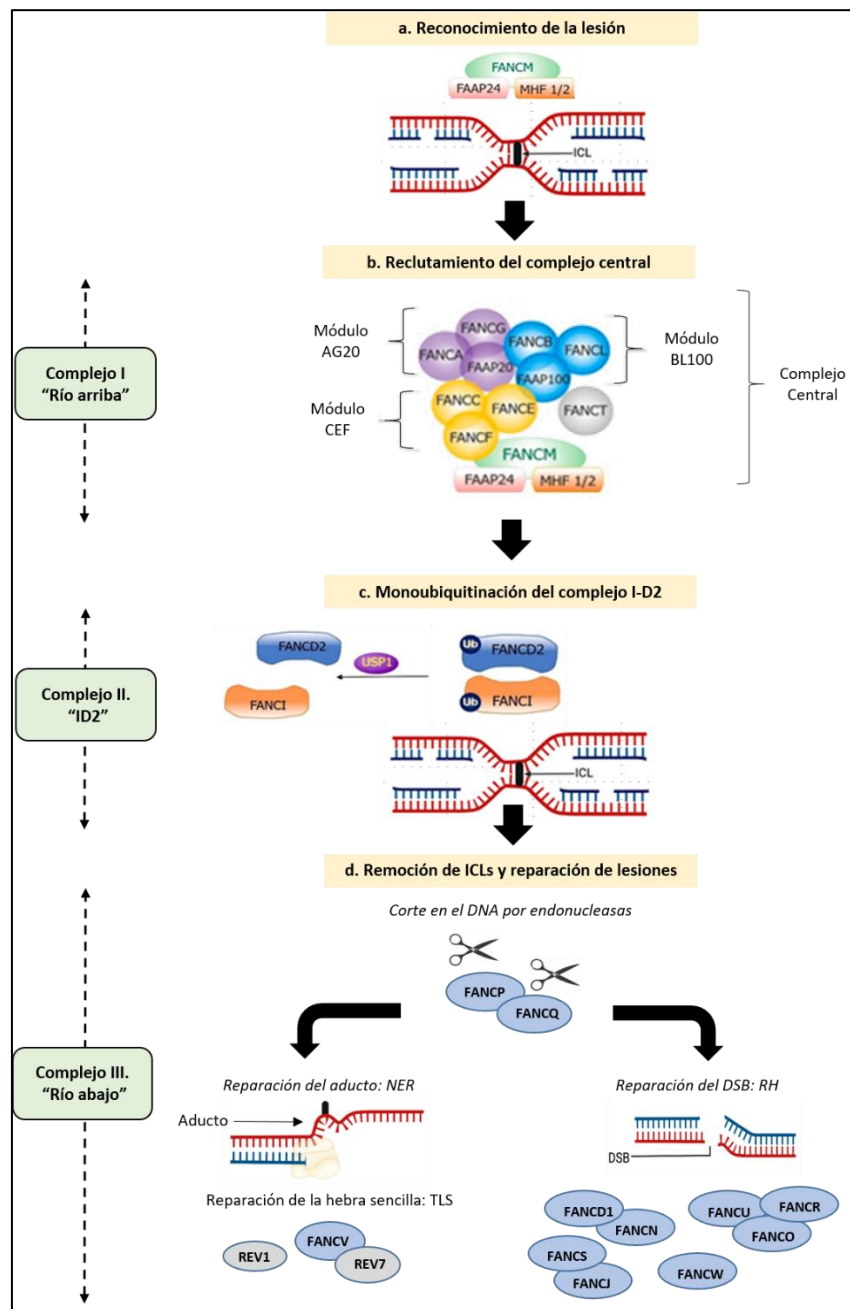


Figura 2. Remoción de ICLs a través de la vía FA/BRCA. a. Reconocimiento de la lesión: FANCM forma parte del complejo ancla con sus proteínas asociadas, y reconoce la convergencia de horquillas de replicación debido a la presencia de un ICL. b. Reclutamiento del complejo central: el complejo central, formado por tres "subcomplejos" (AG20, BL100 y CEF), representa el componente "río arriba" de la vía, y se localiza en el sitio de la lesión. c. Monoubiquitinación del complejo FANCI-FANCD2: El complejo central, por medio de su actividad ubiquitina-ligasa (a través de FANCL), activa al complejo ID2, y este, a su vez, recluta a las proteínas de la vía "río abajo" de la vía. d. Remoción del ICL y reparación de lesiones: luego de la activación del complejo ID2, se reclutan endonucleasas para cortar al DNA y "desenganchar" el ICL, generándose dos intermediarios: el primero contiene un segmento de DNA de una sola hebra y un aducto unido a la hebra opuesta de DNA, que serán reparados por síntesis traslesión (TLS) y reparación por escisión de nucleótidos (NER), respectivamente; y el segundo, que presenta una ruptura de doble hebra (DSB), que será reparada por las proteínas "río abajo", a través de recombinación homóloga (RH).

Además de la participación de las proteínas FANC en la remoción de los ICLs, se ha demostrado que varias de ellas también están involucradas en el mantenimiento del genoma de manera independiente a su función canónica en la ruta FA-BRCA. Por ejemplo, FANCA, FANCB, FANCC, FANCE, FANCG, FANCL, BRCA2, FANCD2 y PALB2 se localizan en el centrosoma y colaboran en el funcionamiento del aparato mitótico (Nalepa et al, 2013). FANCD2 puede interactuar con la helicasa BLM de forma independiente de FANCI, promoviendo el reinicio de las horquillas de replicación bajo condiciones de estrés replicativo (Chaudhury et al, 2013). Asimismo, FANCM también actúa con BLM para prevenir el intercambio de cromátides hermanas y suprimir el alargamiento alternativo de telómeros. Algunas proteínas FANC participan en procesos citoprotectores que no forman parte de la DDR, como la defensa ante especies reactivas de oxígeno, homeostasis mitocondrial y resguardo contra la apoptosis mediada por citocinas pro-inflamatorias (Milletti et al, 2020).

2.6.2 Heterogeneidad genética en la Anemia de Fanconi.

Hasta la fecha, se han identificado 22 genes *FANC* (*FANCA-FANCW*) (Tabla 2.2). La AF resulta de la inactivación de alguno de estos genes, debido a VP germinales.

2.6.2.1 Espectro de VP en los genes *FANC*.

La mayoría de las VP en los genes *FANC* se localizan en *FANCA*, *FANCC* y *FANCG*, abarcando 85%-90% de los casos reportados (García de Teresa et al, 2020).

Las VP encontradas en *FANCA* representan \approx 60%-65% del total reportado en los pacientes, y comprenden variantes de todos los tipos. *FANCA* es un gen formado por 43 exones y las VP descritas suelen ser privadas (es decir, sólo se han vinculado a familias únicas). Por esta razón, el número de variantes reportadas es muy alto, en comparación con el número de pacientes afectados (Castella et al, 2011). Se ha evidenciado que la alta incidencia de deleciones en el gen está asociada a la dispersión de varios elementos repetitivos, como trectos de islas CpG, motivos CCTG/CAGG y secuencias *Alu*. La pérdida de material genómico mediada por las secuencias *Alu* (que pertenecen a los elementos dispersos cortos moderadamente repetitivos) es producto de eventos de recombinación homóloga no alélica entre ellas, debido a su alto grado de identidad y abundancia, especialmente de las subfamilias *AluY* y *AluSx*, que se ubican alrededor del gen y cerca de los sitios de ruptura (Callén et al, 2004; Flynn et al, 2014). Debido a la alta frecuencia de VP en *FANCA*, el cribado molecular en la AF habitualmente inicia con la búsqueda de variantes en este gen, extendiendo luego el análisis hacia otros genes *FANC* (Solomon et al, 2015).

Las VP en *FANCC* comprenden \approx 12%-14% de los casos de AF. Sin embargo, su distribución puede variar dependiendo del grupo étnico, por lo que se ha sugerido que se encuentra entre 5%-14%. *FANCC* tiene 14 exones, y la mayoría de las VP reportadas se concentran entre los exones 1, 4-6 y 13-14. Las dos VP más frecuentes son *FANCC*: c.322delG, asociada a un fenotipo de severidad media, y *FANCC*: c.456+4A>T, encontrada con alta prevalencia en judíos Ashkenazi, así como en población japonesa. Algunas variantes presentes en pacientes con AF también han sido encontradas en casos esporádicos de cáncer pancreático (Gordon y Buchwald, 2013).

Tabla 2. Genes *FANC* identificados hasta la actualidad. Frecuencia y funciones principales en la vía FA-BRCA.

Gen (Alias)	Exones	Amino-Ácidos	Frecuencia	Locus	Funciones principales
<i>FANCA</i>	43	1455	64%	16q24.3	Complejo central, módulo AG20
<i>FANCB</i>	10	859	2%	Xp22.2	Complejo central, módulo BL100
<i>FANCC</i>	15	558	12%	9q22.32	Complejo central, módulo CEF
<i>FANCD1</i> (<i>BRCA2</i>)	27	3418	2%	13q13.1	RH, estabilización de horquillas
<i>FANCD2</i>	43	1471	4%	3p25.3	Complejo ID2, TLS, RH.
<i>FANCE</i>	10	536	1%	6p21.31	Complejo central, módulo CEF
<i>FANCF</i>	1	374	2%	11p14.3	Complejo central, módulo CEF
<i>FANCG</i> (<i>XRCC9</i>)	14	622	8%	9p13.3	Complejo central, módulo AG20
<i>FANCI</i>	38	1328	1%	15q26.1	Complejo ID, monoubiquitinado, TLS, RH.
<i>FANCI</i> (<i>BRIP1/BACH1</i>)	20	1249	2%	17q23.2	Interacción con FANCS, promueve RH e inhibe TLS
<i>FANCL</i> (<i>POG/PHF9</i>)	14	380	0.4%	2p16.1	Complejo central, actividad ubiquitina ligasa (E3)
<i>FANCM</i>	23	2048	0.1%	14q21.2	Complejo central, reconocimiento de ICLs
<i>FANCN</i> (<i>PALB2</i>)	13	1186	0.7%	16q12.2	RH, estabilización de horquillas
<i>FANCO</i> (<i>RAD51C</i>)	9	376	0.1%	17q22	RH
<i>FANCP</i> (<i>SLX4</i>)	15	1834	0.5%	16p13.3	Proteína de andamiaje, regulación de endonucleasas
<i>FANCQ</i> (<i>ERCC4/XPF</i>)	11	916	0.1%	16p13.12	NER, endonucleasa
<i>FANCR</i> (<i>RAD51</i>)	10	340	0.1%	15q15.1	RH
<i>FANCS</i> (<i>BRCA1</i>)	24	1884	0.1	17q21.31	RH
<i>FANCT</i> (<i>UBE2T</i>)	7	197	<0.1%	1q32.1	Complejo central, enzima ubiquitina-conjugante (E2)
<i>FANCU</i> (<i>XRCC2</i>)	3	280	<0.1%	7q36.1	RH
<i>FANCV</i> (<i>MAD2L2/REV7</i>)	11	211	<0.1%	1p36.22	TLS, actividad polimerasa
<i>FANCW</i> (<i>RFWD3</i>)	13	774	<0.1%	16q23.1	RH, actividad ubiquitina transferasa

En esta tabla: Todos los genes *FANC* presentan un patrón de herencia autosómica recesiva, con la excepción de *FANCB*, que es recesiva ligada al X; y *FANCR*, que es autosómica dominante. La frecuencia estimada fue tomada de Shahid et al, 2020. ICL: Enlace covalente cruzado, NER: Reparación por escisión de nucleótidos, RH: Recombinación homóloga, TLS: Síntesis traslesión.

FANCG fue el tercer gen *FANC* en ser identificado. Las VP en este gen representan aproximadamente 8% de los casos en la AF. Está compuesto por 14 exones; y la mayoría de las VP están localizadas en los exones 3, 10, y 12. No se han reportado delecciones grandes. Cerca del 20% son variantes que afectan el proceso de corte y empalme. Se encuentra frecuentemente alterado en pacientes de ascendencia asiática, procedentes de Japón y Corea del Sur (Solanki et al, 2016).

Las variantes en *FANCB*, *FANCD1*, *FANCD2*, *FANCE* y *FANCF*, representan de manera combinada cerca del 13% de las VP reportadas. Las VP en *FANCB* representan 2.2% del total, y constituyen cerca del 4% de los varones reportados en el Registro Internacional de pacientes con AF (abreviado como IFAR en inglés). En *FANCB* no se han reportado variantes de efecto fundador. La naturaleza de sus VP incluye delecciones y duplicaciones largas (20%-30%), además de variantes puntuales e inserciones/delecciones pequeñas; la mayoría conduce a la producción de proteínas truncas. Las alteraciones en este gen se asocian a presentaciones clínicas severas y expectativa de vida baja (Jung et al, 2020). Por otra parte, las variantes bialélicas en *FANCD1* (*BRCA2*) abarcan menos del 5% de los casos con AF, mientras que los hallados en *FANCS* (*BRCA1*) son muy raros. Aunque se han reportado poco más de 50 casos para *FANCD1*, la mayoría desarrolla cáncer de presentación agresiva. Las variantes *FANCD1:c.3492insT* y *FANCD1:c.9424C>T* han sido encontradas en varios casos, y muy pocas veces han sido descritas en estado homocigoto. Respecto a *FANCS*, identificado en pacientes con AF desde 2015, sólo se han reportado 10 casos, de los cuales 6 son homocigotos, afectando principalmente el exón 11 (Woodward y Meyer, 2021). En relación a *FANCE*, este gen posee 10 exones, encontrándose variantes de diversos tipos, siendo prácticamente todas sustituciones de un sólo nucleótido. Las VP en *FANCF*, que está constituido por un sólo exón; representan 2%-3% de los casos con AF reportados, siendo la mayoría puntuales y delecciones pequeñas, asociadas a un fenotipo severo (Favre et al, 2000).

Las VP en los 13 genes *FANC* restantes (*FANCI*, *FANCIJ*, *FANCL*, *FANCM*, *FANCN*, *FANCO*, *FANCP*, *FANCQ*, *FANCR*, *FANCT*, *FANCU*, *FANCV* y *FANCW*) abarcan de manera combinada menos del 5% de los pacientes. Estas VP abarcan cambios puntuales, inserciones y delecciones pequeñas, así como otras que afectan el corte y empalme (Dong et al, 2015). De este grupo, recientemente se reportó una variante de efecto fundador en *FANCL* (*FANCL: c.1021_1092del*) en 13 pacientes de la India y Pakistán, complementando los 9 reportes anteriores (Donovan et al, 2019). En *FANCIJ*, que codifica para una helicasa, destaca la variante *FANCIJ: c.2392C>T*, ubicada en el exón 17, que conduce a un codón de terminación encontrado en 40 de los 58 individuos descritos. Es importante señalar que algunas características clásicamente descritas en la AF, como la FM y la predisposición a leucemia, no han sido reportadas en pacientes con VP en *FANCO*, *FANCR* y *FANCS*, por lo que algunos autores han sugerido que sean considerados como *FA-like*. En este grupo suele incluirse a *FANCM*, ya que el único paciente descrito también tenía VP bialélicas en *FANCA* (Ceccaldi et al, 2016).

2.6.2.2 Variantes patogénicas con efecto fundador.

A pesar de que la AF se considera como una enfermedad rara, la frecuencia de individuos portadores es mucho mayor en ciertas poblaciones. La existencia de grupos étnicos en los que se concentra una gran cantidad de individuos con AF apoya la presencia de un *efecto fundador*, que se presenta cuando en los descendientes de un pequeño grupo de ancestros colonizadores, se amplifican las frecuencias de los alelos afectados, debido a la reducida diversidad genética inicial (Peake y Noguchi, 2022).

Múltiples VP con efecto fundador en los genes *FANC* se han reportado en la literatura, lo cual explica la alta incidencia de la enfermedad en varios grupos étnicos a nivel mundial. Por ejemplo, la frecuencia de portadores de AF en la población judía *Ashkenazi* es de 1 en 89. La variante fundadora en este grupo es *FANCC*: c.456+4A>T (Whitney et al, 1993), la cual muestra un fenotipo severo, con aparición temprana de compromiso hematológico y anomalías físicas mayores. Cabe destacar que pacientes japoneses con el mismo cambio tienen un fenotipo más leve, y esto puede deberse a diferencias en su fondo genético (Futaki et al, 2000).

En población *Afrikáner*, procedente de Sudáfrica, la frecuencia de portadores es de 1 en 77. Las variantes de efecto fundador en este grupo incluyen delecciones del exón 12 al 31, 11 al 17 y la puntual c.3398delA en *FANCA*; entre las tres contabilizan aproximadamente el 80% de los casos con AF en esta etnia. Se piensa que estas VP tienen su origen en los individuos *Afrikáneres* descendientes de los asentamientos de la compañía holandesa del Este de la India, que incluían holandeses, alemanes y franceses protestantes hugonotes, durante el siglo XVII (Feben et al, 2017). Otras poblaciones en las que se han encontrado VP de este tipo incluyen marroquíes (*FANCA*: c.2172-2173insG); alemanes (*FANCG*: c.313G>T), brasileños (*FANCA*: c.3788-3790del) y gitanos españoles (*FANCA*: c.295C>T), siendo este último el que presenta la frecuencia de portadores más alta a nivel mundial, calculada entre 1:64-1:70 (Peake y Noguchi, 2022).

2.7 CARACTERIZACIÓN MOLECULAR DE VARIANTES PATOGENICAS EN LA ANEMIA DE FANCONI.

Una vez que se establece el diagnóstico citogenético de la AF, se debe continuar con la caracterización molecular, a fin de detectar el gen *FANC* afectado, así como el tipo de VP presente en el paciente (Aurbach, 2009). La gran heterogeneidad genética de la AF dificulta la implementación de técnicas de detección clásicamente utilizadas en otro tipo de enfermedades de herencia mendeliana. En el pasado este proceso se lograba a través de la secuenciación Sanger de genes candidatos, frecuentemente realizada luego de estudios de complementación, en los que se sub-tipificaba a los pacientes a través de vectores retrovirales. También era habitual el uso de la cromatografía líquida de alta resolución a partir de zonas codificantes, amplificadas por la reacción en cadena de la polimerasa (PCR), así como la discriminación del complejo de la vía FA-BRCA alterado, al evaluar la monoubiquitinación de FANCD2-FANCI por inmunotransferencia (Ameziane et al, 2008; Bogliolo et al, 2020). No obstante, con el descubrimiento sucesivo de nuevos genes *FANC*, en poco tiempo estas metodologías fueron desplazadas por otras estrategias de genotipificación, con la capacidad para analizar numerosos genes de manera simultánea. Hoy

en día, el estudio de variantes en el DNA de los pacientes incluye la búsqueda de cambios puntuales, deleciones e inserciones pequeñas a través de secuenciación de nueva generación (SNG), así como la exploración de variantes en el número de copias (abreviada como CNV, del inglés *Copy Number Variants*), a fin de identificar deleciones, duplicaciones o inserciones grandes, utilizando diversas técnicas, como la hibridación genómica comparativa, microarreglos de DNA, amplificación de sondas dependiente de ligandos múltiples (MLPA), o análisis de CNV a partir de datos de SNG (Auerbach, 2015; Sroka et al, 2020; Torres et al, 2023).

La SNG ha resultado especialmente útil en la AF para la descripción de variantes en genes *FANC* de frecuencia muy baja, como los descritos más recientemente (Bogliolo et al, 2020). Actualmente, la Guía de Atención Clínica de pacientes con AF señala que la SNG debe utilizarse como parte del proceso de obtención del genotipo, seleccionando un panel de genes en una primera etapa, apelando a la secuenciación de exoma en caso de resultados negativos por panel. La selección del panel puede ser personalizada (elección de genes a estudiar) o a partir de la disponibilidad comercial (por ejemplo, algunos paneles que abarcan diversos síndromes de falla medular, incluyen algunos genes *FANC*) (Sroka et al, 2020). Es importante mencionar que los datos generados a partir de estas plataformas ameritan un análisis bioinformático específico para la búsqueda de variantes, de acuerdo a criterios internacionales (Roy et al, 2018).

Independientemente de la metodología utilizada, una vez que la variante candidata se localiza, el siguiente paso implica la asignación del grado de patogenicidad. El sistema más utilizado para esta tarea proviene de las recomendaciones sugeridas por el Colegio Americano de Genética Médica y la Asociación de Patología Molecular (ACMG-AMP, por sus siglas en inglés). De acuerdo a estos lineamientos, las variantes se clasifican en cinco categorías: benignas, probablemente benignas, variantes de significado incierto (VUS, por sus siglas en inglés), probablemente patogénicas y patogénicas (Richards et al, 2016). En 1998 se creó una base de datos de variantes encontradas en pacientes con AF, como parte de un esfuerzo cooperativo para acelerar el acceso a esta información (*Fanconi Anemia Mutation Database*, <https://www2.rockefeller.edu/fanconi/>).

2.8 APROXIMACIONES A LA CORRELACIÓN GENOTIPO-FENOTIPO EN LA ANEMIA DE FANCONI.

Se ha intentado agrupar diversas características fenotípicas de los pacientes con AF en relación a determinados genes *FANC*, pero este ejercicio ha resultado complejo, debido a que varios de estos rasgos son comunes entre individuos con diverso genotipo. En ocasiones, dicha información está basada en un número limitado de casos, en los que además se pueden reportar datos atípicos con respecto a la presentación clásica de la enfermedad (Neveling et al, 2008). Sin embargo, estas asociaciones se han beneficiado del mejoramiento de las técnicas de genotipificación y de la búsqueda de nuevas formas de realizar la caracterización fenotípica, a partir de un abordaje clínico más detallado. El conocimiento de un gen o una variante específica en la AF es fundamental para el establecimiento de riesgos y opciones terapéuticas (Sroka et al, 2020). Recientemente se publicó una revisión de las posibles correlaciones genotipo-fenotipo físico en la AF, a partir de reportes de casos publicados entre

1982 hasta 2017. En este análisis, en el que se incluyen más 1000 casos, sobresale la falta de descripción de anomalías en el fenotipo en $\approx 50\%$ de los pacientes (Fiesco-Roa et al, 2019).

Sin embargo, de este estudio se extraen varias conclusiones relevantes:

- a. Aquellos casos en los que se reportó al menos una anomalía (de cualquier tipo, $> 80\%$) frecuentemente tenían VP en *FANCB*, *FANCD1*, *FANCD2*, *FANCI*, *FANCI* y *FANCN*. Nótese que, de estos, sólo *FANCB* forma parte de la vía “río arriba”.
- b. Los individuos afectados por VP en *FANCB* presentan con mayor frecuencia, asociación con VACTERL-H ($\approx 80\%$), seguidos de *FANCI* ($\approx 40\%$), en contraste con los casos con genotipo *FANCA* y *FANCG*, que presentan las frecuencias más bajas. Este fenotipo es más común en pacientes con variantes nulas.
- c. Los individuos afectados con VP en *FANCD2* y *FANCI* presentan mayor correspondencia con el acrónimo PHENOS (24% en ambos casos), mientras que este fenotipo es raro en aquellos afectados en *FANCA* y *FANCC*.
- d. Una alta proporción de pacientes con variantes nulas en *FANCA* y *FANCB* presentó al menos una anomalía, en comparación con los casos con variantes hipomórficas.
- e. Pacientes con variantes nulas en estado homocigoto, o hemicigotos, tienen una mayor probabilidad de presentar alguna anomalía, que aquellos con variantes hipomórficas, independientemente de su localización en la vía.

Las correlaciones con las alteraciones hematológicas y la predisposición a cáncer han sido más difíciles de establecer, debido a la discrepancia existente en las observaciones recabadas en diferentes cohortes de pacientes. Por ejemplo, un estudio de casos correlacionó el desarrollo temprano de anemia y leucemias con VP nulas en estado homocigoto en *FANCA*; sin embargo, otros grupos no encontraron dicha asociación (Faivre et al, 2000; Castilla et al, 2011). También se han descrito hallazgos discordantes con respecto a la FM y la supervivencia, cuando se compararon casos con alteraciones en *FANCA* y *FANCG*, con relación a *FANCC* (Faivre et al, 2000; Kutler et al, 2003). Un estudio posterior, realizado en 181 pacientes, concluyó que las malformaciones del eje radial y las anomalías congénitas descritas en el acrónimo CABs (del inglés *congenital abnormality score*) pueden considerarse como factores de riesgo para la evolución de la FM. Dicho riesgo, que fue dividido entre alto y bajo, fue significativamente mayor para pacientes con genotipo *FANCG* y *FANCC*, comparados con *FANCA* (Rosenberg et al, 2003).

Se han encontrado asociaciones más robustas en otros genes, como el riesgo elevado de desarrollo temprano de neoplasias en pacientes con VP en *FANCD1* (*BRCA2*) y *FANCN* (*PALB2*) (Alter et al, 2007; Reid et al, 2007). Por último, es necesario recordar que algunos genes *FANC* son de reciente descubrimiento, y están representados por pocas familias, por lo que se consideran extremadamente raros, de manera que una correlación con el fenotipo podría resultar sesgada (Neveling et al, 2008). Cada vez es más frecuente el estudio de variables más integrativas, como la correlación del fenotipo con la estructura de la vía FA-BRCA, con la localización en regiones exónicas o con VP específicas (Fiesco-Roa et al, 2019; Altintas et al, 2022).

3. PLANTEAMIENTO DEL PROBLEMA.

La presente investigación busca conocer cuál es la distribución de las VP en los genes *FANC* en un grupo de pacientes mexicanos con AF y qué relación tienen con respecto a la presentación de algunas características físicas y hematológicas. Consideramos que este análisis es necesario, debido a que la mayor parte de nuestro conocimiento respecto al espectro genético y las características clínicas en la AF proviene de otras poblaciones, principalmente europeas y occidentales, pero no se ha determinado cual podría ser la representación en el país. De igual manera, se desconoce si las asociaciones establecidas se mantienen o son diferentes a las reportadas, sobre todo considerando la ancestría y diversidad poblacional en México. Por otra parte, los antecedentes señalan que aproximadamente en la mitad de los casos no se ha establecido la correlación genotipo-fenotipo físico en la AF, mientras que su efecto sobre las alteraciones hematológicas requiere de mayor precisión, por lo que la búsqueda de estas asociaciones es relevante para evaluar su potencial sobre la sospecha clínica y su manejo terapéutico.

Asimismo, aunque la pérdida de funcionalidad de la vía FA-BRCA ha sido estudiada de forma amplia en tejido hematopoyético, en especial en modelos *in vitro*, existen pocos reportes que evalúen el efecto del genotipo sobre la progresión de la falla medular y las alteraciones en la proliferación de las CTPH derivadas de los pacientes, motivo por el cual constituyen una oportunidad valiosa para comprender mejor los efectos de las VP sobre la regulación de las células sanguíneas en respuesta al daño genómico.

4. PREGUNTA DE INVESTIGACIÓN.

¿El tipo de VP y el gen *FANC* afectado, están relacionados con las alteraciones del fenotipo clínico y celular, en pacientes con Anemia de Fanconi?

5. JUSTIFICACIÓN.

De acuerdo a la más reciente Guía de Atención Clínica para pacientes con AF, una vez que se obtiene el diagnóstico citogenético, se debe continuar con el abordaje molecular, con el propósito de identificar la variante causal de la enfermedad. La obtención del genotipo en cada caso permite no sólo brindar asesoramiento genético preciso a las familias, sino también identificar posibles donadores de CTPH, además de aportar hallazgos indispensables para integrar correlaciones genotipo-fenotipo.

Existen varios ejemplos en la literatura que describen la distribución de VP en los genes *FANC* en diversas poblaciones, a partir de la cual se saben las frecuencias reportadas a nivel mundial. Desafortunadamente, esta información permanece desconocida en los pacientes con AF en México. Esta situación impide investigar si la heterogeneidad genética tan característica de la enfermedad se replica en nuestra población, o si existen algunas VP predominantes en el país. Esto a su vez representa un obstáculo que impacta negativamente el abordaje terapéutico, considerando las manifestaciones físicas y hematológicas de la AF.

La existencia de pocos reportes de casos en México pone en contexto la necesidad de una estrategia que optimice la inclusión del mayor número de pacientes posible, y sugiere la exploración de alternativas metodológicas centradas en facilitar la búsqueda y análisis de VP en los genes *FANC*. La implementación de estas acciones contribuirá significativamente a comprender el espectro y frecuencia de VP de genes *FANC* en los pacientes con AF en México y de forma paralela, a mejorar el enfoque de diagnóstico, pronóstico y terapéutico de los afectados. La factibilidad del estudio está respaldada por el hecho de que será realizado en un centro de referencia nacional, con una experiencia dilatada en el diagnóstico e investigación tanto clínica como básica en AF. Los resultados de esta investigación podrán ser utilizados para planificar estrategias de genotipificación más eficaces en México, al estar centradas en la casuística derivada del país, así como para enriquecer y profundizar el debate sobre los diversos rasgos fenotípicos encontrado en los pacientes.

6. HIPÓTESIS.

La frecuencia de las VP en los genes *FANC* para la población mexicana, no difiere de las frecuencias reportadas en la literatura. El fenotipo clínico de los pacientes AF, así como el fenotipo celular de sus CTPH, varían dependiendo del tipo y localización de las variantes encontradas.

7. OBJETIVOS.

7.1 General

Identificar y caracterizar variantes patogénicas en los genes *FANC*, en un grupo de pacientes mexicanos con Anemia de Fanconi y realizar una correlación genotipo-fenotipo.

7.2 Específicos.

1. Determinar el tipo y localización de las variantes patogénicas en los genes de la vía FA/BRCA en pacientes mexicanos que acudan al INP en un lapso de 4 años (junio de 2018-junio de 2022).
2. Caracterizar la patogenicidad de las variantes encontradas sobre sus productos proteicos mediante análisis bioinformático.
3. Correlacionar el tipo y localización de las variantes patogénicas encontradas con las características físicas y hematológicas de los pacientes.
4. Correlacionar el tipo y localización de las variantes encontradas con la cantidad y nivel de daño en el DNA en células CD34+.

8. METODOLOGÍA.

8.1 Diseño del estudio.

- a) Finalidad del estudio: Descriptivo.
- b) Secuencia temporal: Transversal.
- c) Control de asignación de los hechos: Observacional.
- d) Inicio del estudio en relación con la cronología de los hechos: Prospectivo.

8.2 Población de estudio.

Pacientes en edad pediátrica con diagnóstico de AF que hayan decidido a participar en el estudio, desde octubre de 2018 hasta octubre de 2022.

8.2.1 Criterios de inclusión.

-Individuos de nacionalidad mexicana, con diagnóstico clínico y citogenético (realizado mediante la inducción de aberraciones cromosómicas con DEB) de AF, de cualquier sexo.

-Firma de carta de consentimiento informado para participar en el estudio.

8.2.2 Criterios de exclusión.

-Pacientes que hubiesen recibido transfusiones durante los tres meses previos al momento de la solicitud de participación en el estudio.

-Pacientes que hubiesen decidido retirarse del estudio.

8.3 Diseño experimental.

Durante el desarrollo del proyecto, cada uno de los pacientes debía llenar una hoja de captación de datos clínicos, y se firmaron las cartas de consentimiento aprobadas por el Comité de Ética del Instituto Nacional de Pediatría.

8.3.1 Estrategia de genotipificación.

Luego de la obtención de las muestras de DNA de los pacientes, a partir de 2-3 ml de sangre periférica, se procedió a la búsqueda de deleciones grandes en *FANCA*, a través de la técnica de amplificación de sondas múltiple dependiente de ligación (MLPA), debido a que

este tipo de variante se ha reportado en una proporción elevada en este gen (25%-30 %), y puede involucrar la pérdida de uno, varios exones, o el gen completo.

Posteriormente, para pacientes que no tuviesen deleciones en *FANCA*, o que se hubiesen encontrado en un sólo alelo, continuamos la genotipificación a través de SNG, utilizando un panel dirigido, conformado por los 16 genes *FANC* más frecuentemente afectados, a fin de detectar VP puntuales o pequeñas.

En caso de que el panel resultase negativo, se continuó la genotipificación por SNG de exoma completo. Si aún no se obtuviese el genotipo por esta vía, se buscaron CNVs a través de MLPA en los genes *FANCB*, *FANCD2* y *FANCN*. Llegado a este punto se esperaba obtener el genotipo de la mayoría de los pacientes estudiados (ver Figura 3).

Si todavía hubiese pacientes sin genotipificar, el método de tamizaje seleccionado fueron los microarreglos de alta resolución, que constan de marcadores a lo largo del genoma para polimorfismos de un sólo nucleótido y para CNVs, a fin de buscar grandes deleciones o duplicaciones en los genes *FANC*, así como detectar tramos largos contiguos de homocigocidad, indicativos de la presencia de alelos homocigotos ininterrumpidos con número de copias neutro (set diploide).

Una vez finalizado el proceso de genotipificación, se procedió a la búsqueda de posibles correlaciones genotipo-fenotipo, partiendo de los datos proporcionados durante la captación de datos y su historia clínica.

8.3.2 Estudio de células troncales y progenitoras hematopoyéticas (CTPH).

En aquellos pacientes en los que fue posible obtener muestras de médula ósea y sangre periférica de manera simultánea, se aislaron células mononucleares y se seleccionaron las CTPH a través del marcador de superficie CD34, a fin de evaluar su abundancia en ambos tejidos. El daño genómico en esta población celular se corroboró a través de la evaluación de la variante de histonas γ H2AX, ya que es indicador de daño ocasionado por rupturas de doble hebra. La detección de los marcajes mencionados se realizó a través de citometría de flujo.

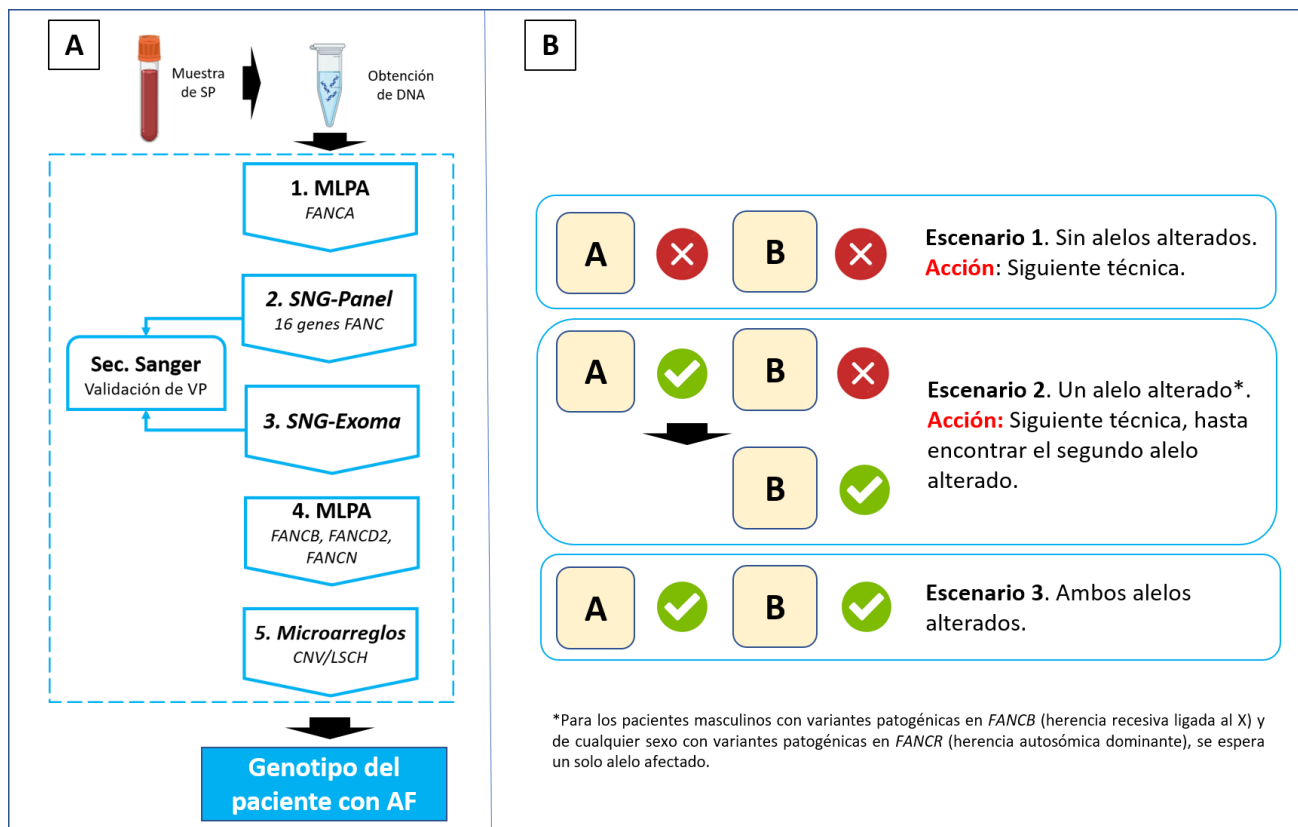


Figura 3. Algoritmo para el diagnóstico molecular del paciente con Anemia de Fanconi (AF): A. Secuencia de metodologías utilizadas para obtener el genotipo del paciente con Anemia de Fanconi (AF): luego de la obtención de DNA genómico a partir de una muestra de sangre periférica, la genotipificación inicia utilizando la técnica de amplificación de sondas múltiple dependiente de ligación (MLPA) para el gen *FANCA*. Posteriormente; se emplean estrategias de secuenciación de nueva generación (SNG), a través del uso de un panel dirigido y de exoma, para la búsqueda de variantes puntuales o deleciones pequeñas en los genes *FANC*. Si aún no se ha determinado la(s) variante(s) responsable(s), el análisis prosigue con el uso de MLPA para los genes *FANCB*, *FANCD2* y *FANCN*; y posteriormente mediante microarreglos de alta densidad, para la búsqueda de variantes en el número de copias (CNV) y de tramos largos y contiguos de homocigocidad (LCSH). B. Criterios de progresión durante el proceso de genotipificación: para cada una de las técnicas utilizadas, existen tres resultados posibles (escenarios): que no se identifique ninguno de los alelos alterados (definidos en la figura como A o B) (escenario 1); que los alelos alterados se identifiquen por técnicas separadas (escenario 2, con la excepción de *FANCB* o *FANCR*), o que las alteraciones en ambos alelos se determinen con la misma técnica (escenario 3).

8.4 Técnicas empleadas.

8.4.1. Obtención de DNA genómico.

El DNA genómico (DNAg) de cada paciente se obtuvo a partir de 3 ml de sangre periférica con anticoagulante (EDTA), utilizando la metodología sugerida por el estuche Gentra Puregene Kit (Qiagen, Venlo, Limburg, NL, USA), con pequeñas modificaciones. Brevemente, un volumen de muestra se transfiere a un tubo cónico de microcentrífuga de 15 ml, y se incuba con 3 volúmenes de solución de lisis de rojos, durante 10 minutos a 37°C. Posteriormente, este tubo se centrifuga a 2000 fuerza centrífuga relativa (FCR) durante 10 minutos, se descarta el sobrenadante y se obtiene el botón de células mononucleares (CMN). En caso de que la muestra conserve glóbulos rojos, se repite el paso de lisis una vez más, respetando la proporción de volúmenes. Al botón de CMN se le añaden 2 ml de solución de lisis de glóbulos blancos, luego se aplica vórtex durante 20 segundos, a temperatura ambiente. Esta mezcla se incuba con 15 µL de RNAsa A (4 mg/ml) a 37°C durante 30 minutos, se enfría en hielo durante 5 minutos y luego se adicionan 15 µL de Proteinasa K (20 mg/ml), a 55°C durante 30 minutos. Posteriormente, el tubo se enfría durante 5 minutos en hielo, y luego se añade 1 ml de solución precipitadora de proteínas, se incuba en hielo por 15 minutos y se centrifuga a 2000 x FCR durante 10 minutos a 4°C. El sobrenadante obtenido se transfiere a un tubo cónico, al que se agrega el mismo volumen de isopropanol. Al agitar esta mezcla por inversión (50 veces, aproximadamente), se observa el DNAg como un hilo blanquecino. Para separar el DNAg, el tubo se centrifuga a 2000 x FCR durante 5 minutos, se remueve el sobrenadante y se conservan aproximadamente 300 µL, que se transfieren junto con el sedimento a un tubo de microcentrífuga de 1.5 ml. El DNAg se lava dos veces con etanol al 70%, para luego dejarlo secar en un termobloque a 55°C, hasta eliminar el líquido remanente. Este botón se resuspende en 50 µL de amortiguador Tris-EDTA de baja concentración (Tris-HCl 10 mM (pH 8.0) + EDTA 0.1 mM) al menos durante 24 horas.

La cantidad y pureza del DNAg se evaluó en un espectrofotómetro Nanodrop TM ND1000 (Nanodrop Technologies, Wilmington, DE). Las diluciones requeridas para los ensayos de SNG, se cuantificaron en un fluorómetro Qubit 4 TM (Thermo Fisher Scientific, EE. UU). La integridad de las muestras se evaluó mediante electroforesis en gel de agarosa estándar al 1%, teñido con Bromuro de Etidio, en amortiguador Tris-Acetato-EDTA 1X.

8.4.2. Técnicas para la detección de variantes.

8.4.2.1 Amplificación de sondas múltiple dependiente de ligación (MLPA).

Para esta técnica se siguió el protocolo sugerido por el fabricante (MLPA General Protocol (*One tube*), MDP-v005, MRC Holland). Se utilizaron las mezclas de sondas SALSA correspondientes para la detección de CNVs para cada gen de interés (ver Tabla 8.1). Antes de comenzar, las muestras de DNAg en cada ensayo se diluyeron hasta una concentración final de 16ng/µL-18ng/µL, para obtener 80ng-90ng totales en 5 µL. En cada corrida se incorporaron entre 8-10 muestras de pacientes, y se adicionaron 4 muestras de referencia (DNA masculino y DNA femenino de individuos sanos, mezcla sin DNA y control positivo) (Ver Tabla 8.1).

Tabla 3. Características de los estuches de MLPA.

Gen evaluado	Mezcla de sondas	Versión de lote
<i>FANCA</i>	P031/P032	B2, B3
<i>FANCB</i>	P113	B1
<i>FANCD2-FANCN</i>	P057	B2

La metodología de MLPA consta básicamente de 6 pasos. Brevemente, las muestras de DNAg (5 µl) se desnaturalizan durante 5 minutos a 98°C. Luego, se agregan 3 µl de la mezcla de sondas a hibridar (1,5 µl de mezcla de sondas SALSA + 1,5 µl del tampón amortiguador de MLPA) por muestra de DNA. Después, los tubos se colocan durante 1 minuto a 95°C y se incuban durante 16 horas a 60°C. Transcurrida la hibridación, se disminuye la temperatura a 54°C y se agregan 32 µl de la mezcla de ligación, a la misma temperatura durante 15 min. Para inactivar la Ligasa 65, la mezcla se calienta durante 5 minutos a 98 °C. Posteriormente, los tubos se enfrían a temperatura ambiente y se añaden 10 µl de mezcla de polimerasa. La reacción en cadena de la polimerasa (PCR) se inicia utilizando el programa sugerido (95°C por 30 segundos, 60°C por 30 segundos y 72°C por 1 minuto, durante 35 ciclos), con una amplificación final a 72°C por 20 minutos. Los tubos se protegen de la luz hasta la electroforesis capilar.

Para discriminar el tamaño de los productos de PCR, se utilizó un instrumento de electroforesis capilar Applied Biosystems 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), agregando como marcador de peso molecular el *GeneScan*™ 500 LIZ. Los archivos de esta separación se analizaron con el programa Coffalyzer (MRC Holland, <https://www.mrcholland.com/technology/software/coffalyser-net>), siguiendo las sugerencias del programa para la determinación del número de copias y los valores de coeficiente de dosis, obtenidos luego de la normalización y contraste de las señales de las muestras de referencia y las muestras de los pacientes, para cada sonda evaluada (ver Figura 4).

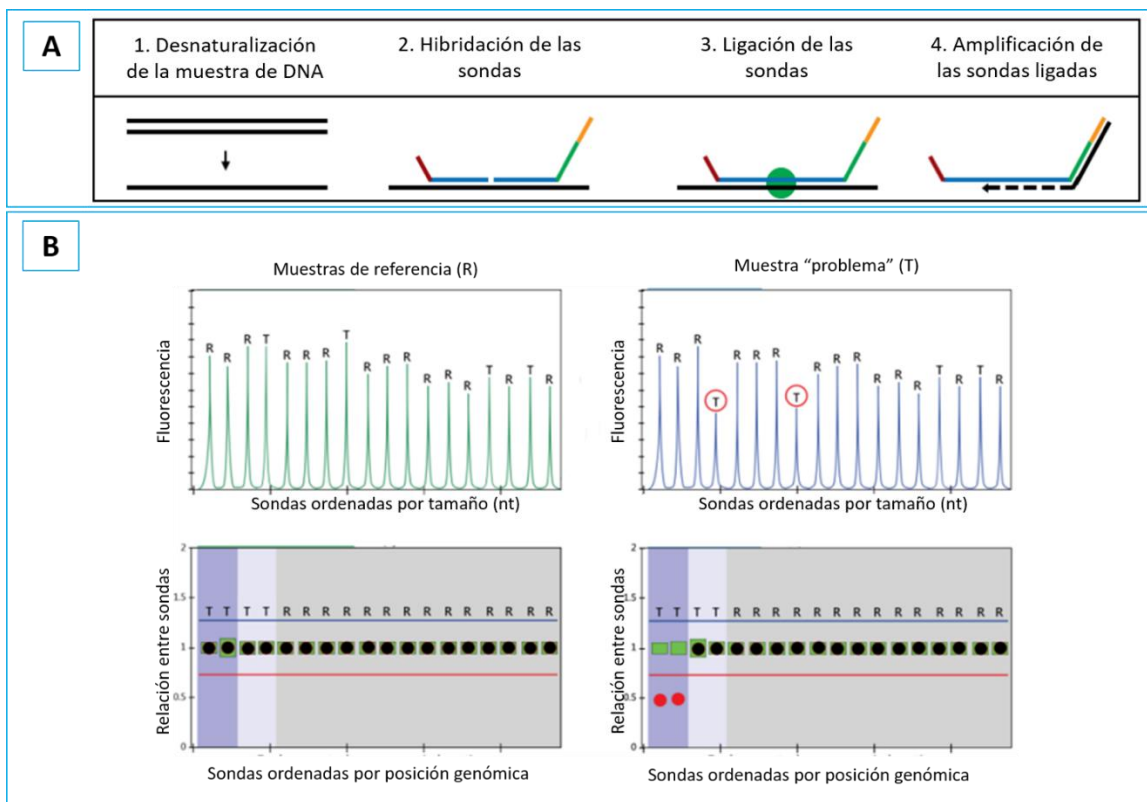


Figura 4. Descripción de la técnica de MLPA para la detección de CNVs. A. Durante la fase inicial, la muestra de DNA se desnatura, y se mezcla con un conjunto de sondas específicas (paso de hibridación). Cada sonda de MLPA consta de dos oligonucleótidos (paso 2), los cuales deben hibridarse con secuencias diana contiguas para poder ser ligados entre sí y formar una sonda única (paso 3). Posteriormente, durante la PCR, todas las sondas ligadas son amplificadas simultáneamente usando el mismo par de cebadores, que tienen acoplado un fluorocromo (paso 4) B. Electroforesis capilar y cuantificación relativa: Al comparar la altura relativa de cada uno de los picos generados por las sondas (proporcional a la intensidad de fluorescencia) entre una muestra problema (parte derecha de la figura) con los de las muestras de referencia (parte izquierda de la figura, sección superior), se observa una disminución de dos sondas en la muestra problema (círculos rojos). En la parte inferior de la figura se visualizan las relaciones de las sondas de la muestra problema (parte derecha) normalizadas con las muestras de referencia (parte izquierda), y ordenadas de acuerdo a su localización cromosómica (en el programa Coffalyzer.net).

8.4.2.2 Secuenciación de nueva generación (SNG).

8.4.2.2.1 Panel de SNG.

El diseño *in silico* de los oligonucleótidos específicos para 16 genes *FANC* (*FANCA*, *FANCB*, *FANCC*, *FANCD1/BRCA2*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *FANCI/BRIP1*, *FANCL*, *FANCM*, *FANCN/PALB2*, *FANCO/RAD51C*, *FANCP/SLX4* y *FANCO/ERCC4*) se realizó con la aplicación web SureDesign (Agilent Technologies). Se establecieron como parámetros la cobertura del 99,5% de los exones codificantes de estos genes más 50 pb de sus regiones intrónicas 5' y 3' flanqueantes. Esta plataforma es compatible con el sistema de enriquecimiento HaloPlex (Agilent Technologies Inc., 2013, Waghäusel-Wiesental, Alemania),

que se utilizó para la digestión del DNA (a partir de 225 ng), la síntesis de bibliotecas y la amplificación por PCR de los fragmentos seleccionados. La SNG se llevó a cabo de acuerdo al método de *secuenciación por síntesis con terminadores reversibles* desarrollado por Illumina, obteniéndose fragmentos de 150 pb leídos desde ambos extremos (*paired-end sequencing*) en un equipo MiSeq (Illumina Inc, San Diego, CA, USA).

8.4.2.2.2 Secuenciación de exones (Exoma).

Para esta etapa de SNG, las muestras fueron enviadas a un servicio externo (Novogene Co, LTD, Beijing, China). La amplificación de las regiones exónicas se realizó con el estuche SureSelect Human All Exon + UTR, V6 (Agilent Technologies Inc., 2013, Waghäusel-Wiesental, Alemania), con el que se capturan ~50 Mb del genoma humano, a partir de 200 ng de DNA. El método de secuenciación utilizado también se fundamenta en el mismo principio (*secuenciación por síntesis, paired-end*) descrito en la sección anterior, en un equipo HiSeq 2000 (Illumina Inc, San Diego, CA, USA).

8.4.2.2.3 Análisis bioinformático.

El flujo bioinformático (*pipeline*) para el llamado de variantes se realizó de acuerdo a la guía de buenas prácticas recomendadas por GATK (*Genome Analysis Toolkit*) (DePristo et al., 2011; Van der Auwera et al., 2013). Brevemente, los archivos recibidos (en formato .fastq) se procesaron en un sistema operativo Linux, de acuerdo a los pasos que se señalan a continuación (los programas empleados se señalan en paréntesis):

- a. Revisión de la calidad de las lecturas de secuencia (FastQC).
- b. Filtrado de bases de mala calidad y remoción de adaptadores (Trimmomatic).
- c. Alineamiento y mapeo de las lecturas filtradas (alineación de Burrows-Wheeler (BWA-MEM), usando las referencias para genoma humano que se encuentran en bases de datos como UCSC o Ensembl).
- d. Transformación de archivos SAM en formato BAM y marcaje de duplicados (Samtools y Picard).
- e. Inspección de la calidad del mapeo (revisión de archivos BAM) (Qualimap).
- f. Obtención del archivo de llamado de variantes (VCF), en modo gVCF (genotipificación conjunta).
- g. Anotación de variantes (VEP y SnpEff). Las variantes se describieron de acuerdo a las pautas de nomenclatura de la *Human Genome Variation Society* (HGVS) (den Dunnen et al., 2016).

En el caso de la lecturas derivadas del panel, las regiones codificantes fueron filtradas a través de un archivo .BED (del inglés *Browser Extensible Data*), generado a partir de las coordenadas de los genes *FANC* en la versión hg19 del genoma humano, en la plataforma UCSC Genome Browser (<https://genome.ucsc.edu/cgi-bin/hgTables>).

8.4.2.2.4 Validación de variantes.

En los pacientes en los que se encontró una variante con potencial patogénico por SNG, ya sea por panel o mediante exoma, estas se validaron a través de secuenciación de Sanger. Para ello, se diseñaron oligonucleótidos de 18-23 nucleótidos, para la síntesis de productos de PCR punto final de 250-300 pb. Las condiciones de reacción, así como el programa de amplificación, se muestran en la Tabla 4 y Tabla 5. Cada producto se corroboró por electroforesis en gel de agarosa al 2%. Los productos se purificaron con el estuche Qiaquick PCR Purification Kit (QIAGEN, Venlo, Limburg, NL, EE. UU.), siguiendo las instrucciones del fabricante. Para la secuenciación, se utilizó el estuche Big Dye Terminator Kit, en un equipo ABI Genetic Analyzer 3130 (Applied BioSystems, Foster City, EE.UU.). Los electroferogramas se analizaron en el programa Chromas v2.6.6 (www.technelysium.com.au), comparando las secuencias obtenidas con las de referencia, a través de la herramienta web BLAST/BLAT.

Tabla 4. Condiciones de reacción para PCR.

Reactivos	Concentración inicial	Concentración final	Volumen final
Agua libre de nucleasas	-	-	Hasta 25 μ L
Tampón de reacción	10X	1X	2.5
Cloruro de Magnesio	25 mM	1.5 mM	1.5
Mezcla de dinucleótidos	10 mM c/u	200 μ M	0.5
Oligo "Sentido"	10 mM	0.25 μ M	0.625
Oligo "Anti-sentido"	10 mM	0.25 μ M	0.625
Enzima Taq Polimerasa	5U/ μ L	0.75U	0.1875
DNA genómico		120-150 ng	

Tabla 5. Programa de amplificación.

Pasos	Temperaturas		Tiempo
Desnaturalización	95°C		5 minutos
Fase de PCR	95°C	35 ciclos	25 segundos
	55°C-60°C (*)		30 segundos
	72°C		30 segundos (*)
Extensión final	72°C		7 minutos
Mantenimiento	4°C		∞

(*) Dependiente de las condiciones de cada amplicón.

8.4.2.3 Secuenciación de Sanger dirigida.

En caso de sospechar que la variante causal de la AF en los pacientes estuviese asociada a un efecto fundador previamente reportado, el abordaje molecular se realizó por Secuenciación de Sanger dirigida. La metodología fue la misma que la descrita en la sección 8.4.2.2.4, considerando la localización geográfica e historia demográfica de los grupos étnicos a los que pertenecían los pacientes con AF estudiados.

8.4.2.4 Uso de microarreglos de alta densidad.

Para el estudio de DNAG mediante microarreglos, se utilizaron los soportes CytoScan750K (Affymetrix, Santa Clara, CA, EE. UU.); que proporcionan cobertura de todo el genoma, e incluyen 550,000 marcadores para la detección de CNVs y 200,436 sondas para SNVs. La digestión, ligación, amplificación por PCR, fragmentación, etiquetado, desnaturalización e hibridación del DNA en el microarreglo, se realizó de acuerdo a la metodología proporcionada por el fabricante.

Luego, los microarreglos se tiñeron y lavaron en la estación de líquidos Affymetrix GeneChip Fluidics Station 450, y se escanearon en un equipo Affymetrix GeneChip Scanner 3000 7G (Affymetrix, Santa Clara, CA, EE. UU.). Los archivos obtenidos se analizaron con el programa Chromosome Analysis Suite (ChAS), versión 4.1 (Affymetrix, Santa Clara, CA, EE. UU.), utilizando como archivos de anotación los proporcionados en la versión NetAffx 33. Para el análisis de CNVs, se utilizaron parámetros de “alta resolución” (ganancia y pérdidas en 100Kb con ≥ 50 marcadores (sondas)).

8.4.2.4.1 Determinación de segmentos largos y contiguos de homocigocidad (LCSH).

Los umbrales mínimos para los LCSH en análisis clínicos generalmente se establecen entre 3 y 10 Mb (Kearney et al., 2011). Los LCSH se visualizaron en el programa ChAS, versión 4.1. Para LCSH >3 Mb, el análisis se configuró en LOH con recuento de marcadores (sondas) = 50, y para LCSH >5 Mb se estableció en recuento de marcadores = 50. En caso de sospecharse de la presencia de un haplotipo dentro de una LCSH compartida entre pacientes, se seleccionaron y filtraron los marcadores considerando su configuración alélica, para usarlos en el módulo LDhap de la plataforma LDlink (<http://analysistools.nci.nih.gov/LDlink/>), a fin de calcular la frecuencia de haplotipos específicos.

8.4.2.4.2 Determinación global de ancestría.

Este análisis se realizó en aquellos casos en los que se precisaba indagar sobre el posible origen étnico de pacientes con una variante previamente no reportada, en los que se sospechó la presencia de un efecto fundador. Para ello, los archivos CEL procedentes de los microarreglos fueron estudiados en el programa ChAS (versión 4.1), para obtener la configuración de genotipos para cada marcador de SNV, a través del algoritmo BRLMM-P-plus. Para contrastar esta información con datos de microarreglos de alta densidad de Affymetrix ubicados en el repositorio del Proyecto de los 1000 Genomas (1KGP, <https://www.internationalgenome.org/>), se utilizó el programa PLINK, versión 1.9 (<https://zzz.bwh.harvard.edu/plink/download.shtml/>). Los SNVs no coincidentes o

discordantes fueron filtrados y removidos. Finalmente, la inferencia de la ancestría global de la población mexicana se realizó a través de un enfoque de máxima verosimilitud, con el programa ADMIXTURE (Alexander et al, 2009), a partir de 3 componentes ancestrales. Para este análisis, las frecuencias genotípicas de las poblaciones de Europa del Norte de Utah (CEU), Yoruba (YRI) y de Nativos Americanos (NAT), reportadas en el 1KGP, se consideraron como parentales de referencia. Para minimizar el efecto de la estratificación poblacional, se consideraron 345 variantes de nucleótido único, previamente descritas como marcadores informativos de ancestría (Barbosa et al, 2017).

8.5 Determinación de patogenicidad de variantes candidatas.

La clasificación de las variantes candidatas se realizó de acuerdo a los criterios sugeridos por la guía de la ACMG-AMP (Richards et al, 2015). Los criterios para la clasificación de variantes en esta guía permiten agrupar a las mismas en cinco clases: benignas, probablemente benignas, variantes de significado incierto, probablemente patogénicas y patogénicas. Para realizar este ejercicio se utilizaron dos herramientas integrativas en línea: Varsome (<https://varsome.com/>) y Franklin Genoox (<https://franklin.genoox.com/clinical-db/home>). A través de ellas se pueden consultar datos importantes como frecuencias alélicas por población para un determinado cambio, si dicha variante ha sido reportada previamente, entre otras. Adicionalmente, se analizaron datos derivados de los expedientes clínicos, así como de programas bioinformáticos (ver siguiente sección), para establecer un veredicto final sobre el efecto deletéreo de cada una de las variantes evaluadas.

8.5.1 Análisis bioinformático de patogenicidad.

Para la predicción del impacto funcional de una variante no sinónima (de sentido erróneo) a nivel de región codificante, se utilizaron los programas SIFT 4.0, Polyphen 2, MutPred2 y SNAP2. Para evaluar dicho efecto sobre la estabilidad de las proteínas, se utilizaron los programas MUpro y I-Mutant 3.0. Adicionalmente, para estudiar la conservación filogenética de la posición, se emplearon los programas ConSurf y PHANTHER. La revisión de los cambios a nivel de interacciones estructurales entre aminoácidos, se visualizó con el programa PyMOL Stereo 3D, a partir de archivos .pdb extraídos de AlphaFold y analizados en Dynamut2. Para el caso de las variantes que conducen a corrimiento de marco de lectura o a un codón de paro, se utilizaron los programas MutpredLOF y VEST-indel. En el caso de variantes asociadas al proceso de corte y empalme, se utilizaron cuatro herramientas para evaluar su patogenicidad: Human Splicing Finder, MaxEntScan, RNABPS y CRYP-SKIP. La información asociada a los algoritmos utilizados por los programas mencionados se encuentra en la tabla suplementaria S1.

8.6 Recopilación de datos clínicos de los pacientes con Anemia de Fanconi.

8.6.1. Características físicas.

Durante la revisión de los pacientes con AF se usaron datos clínicos y paraclínicos para su descripción. Las características físicas se agruparon en 2 categorías. Se consideraron las anomalías presentes en la asociación VACTERL-H (**V**ertebrales, **A**nales, **C**ardíacas, fístula **T**raqueo-esofágica, atresia **E**sofágica-duodenal, **R**enales, en las extremidades (**L**imbs, en inglés) e **H**idrocefalia), así como las propuestas en el acrónimo PHENOS (Alter y Giri, 2016), que engloban alteraciones en la **P**igmentación, cabeza pequeña (**S**mall **H**ead, en inglés), alteraciones **N**euroológicas, **O**tológicas y estatura baja (**s**hort **S**tature, en inglés). Se considera que un paciente presenta la asociación VACTERL-H si presenta al menos 3 de las 8 características, mientras que para PHENOS, debe presentar al menos 4 de las 6 características.

8.6.2. Características hematológicas y oncológicas.

La severidad de la falla medular (FM) se definió en tres categorías, de acuerdo al grado de citopenia observada, siempre y cuando fuese persistente y no secundaria a otra causa tratable, según los parámetros descritos en la Guía de manejo de pacientes con AF (Sroka et al, 2022), a partir de los datos de biometrías hemáticas obtenidas en la fecha más próxima al diagnóstico de la AF. La presencia de síndrome mielodisplásico (SMD), de alguna neoplasia hematológica, o de tumores sólidos, se obtuvo a partir del diagnóstico de estas condiciones, reportadas en los expedientes médicos.

8.7. Determinación estadística de las correlaciones.

Los pacientes fueron agrupados de acuerdo a su genotipo, ubicación del gen *FANC* alterado en la vía FA-BRCA y al tipo de variante. En el último caso, las que conducen a interrupción del marco de lectura, codón de paro, pérdida del inicio, deleciones ≥ 4 nucleótidos o aquellas que afecten sitios canónicos del proceso de corte y empalme se clasificaron como nulas, mientras que las variantes de sentido erróneo, deleciones o inserciones dentro del marco de lectura o aquellas de corte y empalme, cuya predicción causaba cambios en la estructura de la proteína, fueron clasificadas como hipomórficas.

La relación entre las características fenotípicas y la información genética encontrada, se visualizó utilizando los paquetes Seaborn (Python 3.9) y ggplot2 (RStudio, Boston, MA). Se utilizó la prueba exacta de Fisher para la búsqueda de asociaciones entre variables cualitativas (en este caso, comparaciones entre la frecuencia de los rasgos valorados), utilizando un valor de $p < 0.05$.

8.8 Análisis de células troncales y progenitoras hematopoyéticas de pacientes con Anemia de Fanconi a través de citometría de flujo.

8.8.1 Separación de células mononucleares (CMN) y marcaje.

Este aislamiento se realizó a partir de muestras pareadas de sangre periférica y médula ósea de pacientes con AF, a partir de un volumen inicial de 2 ml. Primero, se lisaron los glóbulos rojos de acuerdo a lo descrito en la sección 8.4.1, hasta obtener CMN. Posteriormente estas células se lavaron con 1 ml de PBS 1X, para luego contabilizar y evaluar su viabilidad a través del ensayo de exclusión con Azul Tripano al 0.4%, en un hemocitómetro (Marien Field, Alemania). Se calculó el volumen necesario para obtener 200,000 células, que fueron resuspendidas en suero fetal bovino (Gibco, BRL, Grand Island, NY, USA) hasta 500 μ l, e incubadas a 4°C durante 15 minutos. Luego se añadieron 500 μ l de tampón fosfato salino (PBS) 1X a cada tubo, y se centrifugaron a 2000 x FCR durante 5 minutos. Se retiró el sobrenadante hasta conservar aproximadamente 100 μ l, en los que se diluyeron 1 μ l de marcador de superficie común de leucocitos CD45 (APC Mouse Anti-human, BD Bioscience, número de catálogo 555485) y 1 μ l de marcador de superficie para células progenitoras hematopoyéticas CD34 (FITC anti-human CD34, Clon 561, Biolegend, número de catálogo 343604), colocándolos a 4°C durante 45 minutos. Transcurrido ese tiempo, se siguió el protocolo sugerido por el estuche Pharmigen TM Apoptosis, DNA Damage and Cell Proliferation Kit (Beckton Dickinson, NY, USA, número de catálogo 562253). Este estuche permite el marcaje de células con daño, a través de la detección de la fosforilación de la variante de histona H2AX, en rupturas de doble hebra en el DNA. Una vez marcadas, las células se almacenaron a 4°C y se protegieron de la luz, por un periodo no mayor a 72 horas. El procedimiento descrito se realizó de igual manera en las muestras procedentes de individuos sanos, que fueron utilizadas para fines comparativos.

8.8.2 Detección de señales fluorescentes y análisis de datos.

La detección de señales de fluorescencia se realizó en un citómetro Attune NxT (Thermo Fisher Scientific, EE.UU.), previa separación en base a su tamaño y complejidad de al menos 150,000 células sanguíneas. Los archivos de captura se analizaron con el programa FlowJo v10 (Beckton Dickinson, USA), de acuerdo a los criterios recomendados para la estimación de la compensación, separación de poblaciones celulares y cálculo de intensidad de fluorescencia media (MFI, por sus siglas en inglés) para cada uno de los fluoróforos utilizados. Una vez obtenidos los resultados finales, las pruebas estadísticas para análisis comparativo (Pruebas de U de Mann-Whitney y Exacta de Fisher, con $\alpha=0.05$) se realizaron en el programa RStudio (RStudio, Boston, MA).

9. RESULTADOS.

9.1 Breve descripción de la población de estudio.

Durante el transcurso del presente proyecto, y habiendo cumplido con los criterios necesarios para su inclusión en el mismo, se capturaron muestras de 58 pacientes. En cada uno de ellos se aplicó al menos una de las metodologías sugeridas en la estrategia de genotipificación. Debido a que el grado de afectación hematológica de los pacientes repercutió sobre la cantidad de DNA que se pudo extraer de las muestras obtenidas, en algunos casos la misma resultó insuficiente para llevar a cabo las técnicas necesarias para la resolución del genotipo. Por esta razón, se obtuvo el genotipo responsable de la AF en 41 casos, pertenecientes a 34 familias. De esta manera, la eficiencia diagnóstica fue de 70.68%. Las características demográficas de los pacientes con genotipo se resumen en la Tabla 6, en la que se observa que la proporción de pacientes femeninas casi duplica a los pacientes masculinos (1.92X). De este grupo, tres pacientes (7.31%) desarrollaron neoplasias: uno leucemia mieloide aguda, uno síndrome mielodisplásico y otro meduloblastoma.

Tabla 6. Características demográficas de los pacientes con genotipo.

Datos	Totales
Número de pacientes	41
Casos únicos vs familiares	33 vs 8
Sexo (M:F)	14:27
Promedio de edad al diagnóstico (rango, en años)	7.53 (0.29-17.07)

En la tabla: F: pacientes femeninos, M: pacientes masculinos.

9.2 Resultados que cubren el objetivo particular 1: Determinar el tipo y localización de variantes patogénicas en los genes de la vía FA/BRCA en pacientes mexicanos que acudan al INP en un lapso de cuatro años (Junio de 2018 a Junio de 2022)

9.2.1 Resultados de acuerdo al gen FANC afectado.

Se encontraron variantes con potencial efecto patogénico en nueve genes FANC. Los afectados en FANCA constituyen el 56.1% de los 41 pacientes con genotipo, comprendiendo 23 pacientes. De acuerdo a la frecuencia obtenida, el segundo gen más afectado fue FANCB, con alteraciones en cuatro pacientes, seguido por FANCC, FANCG y FANCF (tres pacientes por gen). Dos pacientes presentaron VP en FANCE, mientras que las VP en FANCD1, FANCD2 y FANCI están representadas por un sólo paciente cada uno (Figura 5, A).

Respecto a la configuración alélica, más del 60% de los pacientes fueron homocigotos, mientras que 26.83% fueron heterocigotos compuestos (debido a que no se encontraron pacientes con VP en *FANCR*; que se heredan de forma autosómica dominante). (Figura 5, B). Por otra parte, la gran mayoría (>90%) de los pacientes afectados presentó variantes en genes del complejo “río arriba” de la ruta FA-BRCA (*FANCA*, *FANCB*, *FANCC*, *FANCE*, *FANCF* y *FANCG*), en contraste con lo encontrado para los otros dos complejos (“I-D2” y “río abajo”), que en conjunto sólo comprenden tres pacientes (Figura 5, C). Los pacientes con AF y genotipo determinado provenían de 13 estados de México, y se localizaron en mayor número en la región central del país. (Figura 6).

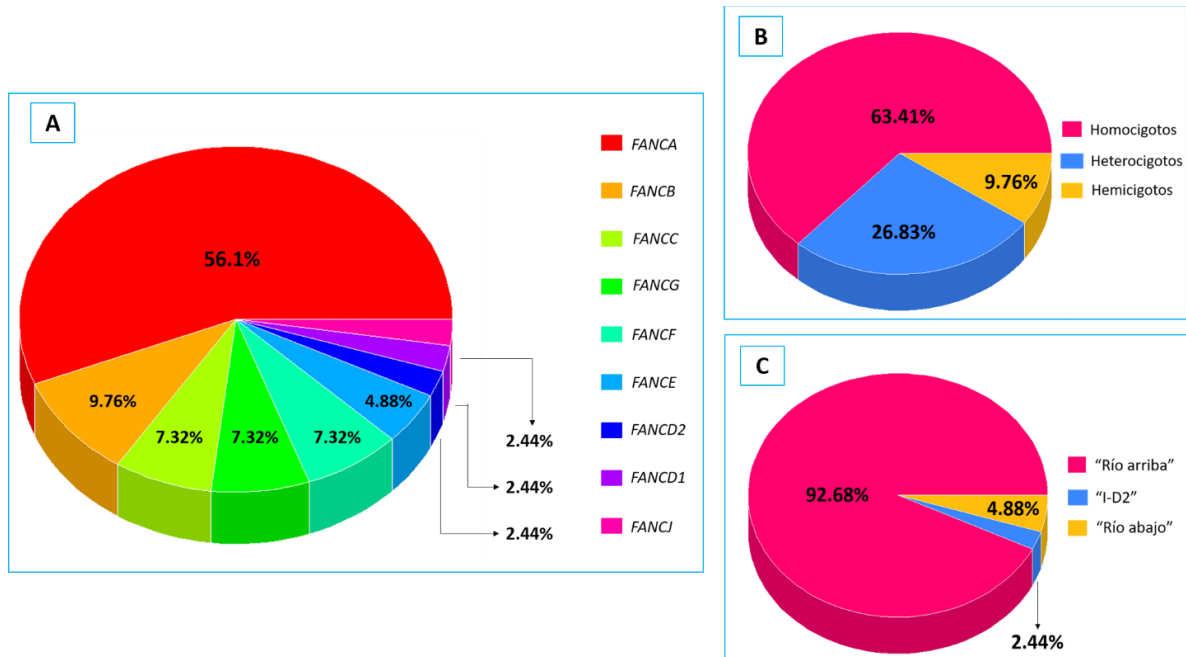


Figura 5: Distribución de 41 pacientes mexicanos con Anemia de Fanconi, de acuerdo al genotipo, cigosidad de las variantes y parte de la vía FA/BRCA afectada. A. Distribución de acuerdo al gen *FANC* afectado. B. Distribución de acuerdo a la configuración alélica de las variantes. C. Distribución de acuerdo a la ubicación en la vía FA/BRCA, considerando los genotipos obtenidos.



Figura 6. Distribución de 41 pacientes mexicanos con Anemia de Fanconi, de acuerdo a su ubicación geográfica. En la figura, el número en el interior del círculo amarillo crema con borde azul representa la cantidad de pacientes con AF con genotipo en ese Estado. Los pacientes procedían de 13 estados: Baja California, Sinaloa, Tamaulipas, Nuevo León, Veracruz, Oaxaca, Guerrero, Ciudad de México, Estado de México, Puebla, Morelos, Hidalgo y Guanajuato.

9.2.2 Resultados de acuerdo al tipo y localización de las variantes.

En los genes *FANC* afectados se encontraron 33 VP únicas. Las más frecuentes fueron las VP no sinónimas (de sentido erróneo), así como aquellas que ocasionan cambios en el marco de lectura, que se identificaron en la misma proporción (21.95%); seguidas por las alteraciones en el corte y empalme (12.2%) y las variantes que provocan un codón de paro (9.76%) (Figura 7, A). En este análisis, destaca el número y diversidad de las variantes presentes en *FANCA* (18 VP), en contraste a lo descrito para el resto de los genotipos (8 genes *FANC*, 15 VP). De este último grupo, sólo en *FANCB* se presentaron más de dos VP. En tres genes (*FANCC*, *FANCG* y *FANCI*) se encontró una sola VP. En relación al estado de reporte, cerca de la mitad de las variantes identificadas no se han descrito previamente (Figura 7, B).

Por otra parte, considerando la estrategia metodológica propuesta para la genotipificación de los pacientes con AF (ver Figura 3), la mayoría de los alelos afectados en los genes *FANC* corresponden a variantes puntuales y deleciones/duplicaciones pequeñas, determinadas a través de métodos de SNG (>60%). Es importante señalar que a través de la metodología de MLPA se detectaron cambios en el número de copias (CNV) sólo en *FANCA* (19.56% del total para ese gen) (Figura 7, C). No se detectaron alteraciones en CNV a través de microarreglos. La secuenciación de Sanger fue utilizada para la confirmación del genotipo en hermanos afectados para casos familiares, cuyas variantes fueron identificadas a partir de SNG, salvo en el caso de *FANCC*, donde fue utilizada de forma dirigida para la detección de una variante de efecto fundador (ver sección 9.3.2).

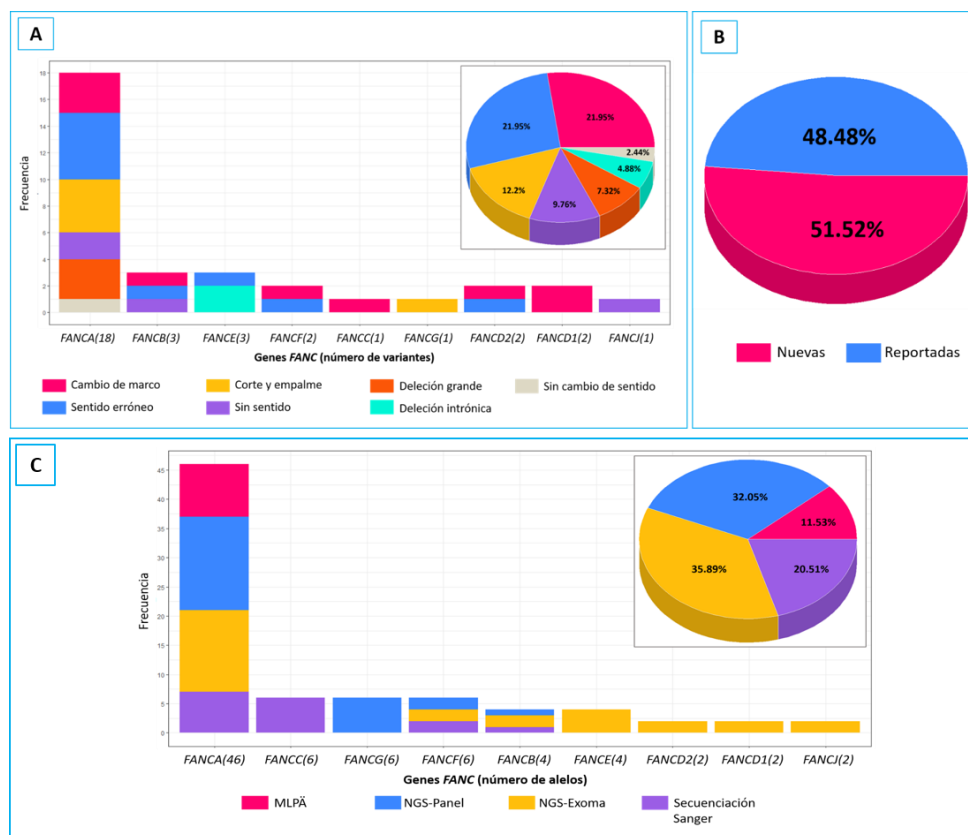


Figura 7. Espectro de variantes encontradas en los genes *FANCA* en este estudio, de acuerdo a su localización, estado de reporte y metodología de identificación. A. Tipo de variante de acuerdo a su localización por gen. En el recuadro superior derecho se señala la proporción de las mismas, con respecto a la totalidad encontrada. B. Estado de reporte de variantes, de acuerdo a la información almacenada en la base de datos *Fanconi Anemia Mutation Database*. C. Proporción de alelos alterados, detectados de acuerdo a las técnicas empleadas; los números en paréntesis al lado del nombre del gen corresponden a la sumatoria total de los alelos. En el recuadro superior de la imagen se señala la proporción de las mismas, con respecto a su totalidad.

De la totalidad de variantes encontradas, 18 se localizaron en *FANCA*. Como se observa en la Figura 8, A; la proteína codificada por este gen posee cinco señales de localización nuclear, además de diversos sitios de interacción con otras proteínas de la vía FA-BRCA. La variante con más alelos afectados (siete) en este estudio fue *FANCA*: c.2535_2536delCT (tres pacientes en estado homocigoto, y un paciente en estado heterocigoto), que ocasiona un cambio de marco de lectura que genera una proteína trunca (p.C846Qfs*20) carente del sitio de unión a FAAP20 (Figura 8, B); seguida por una delección grande, que abarca los exones 1-3 (seis alelos, en tres pacientes en estado homocigoto). Aunque se observa una disposición dispersa de las VP a lo largo del gen, así como la diversidad de las mismas, la mayoría se concentra hacia la mitad amino terminal, que coincide con la región de unión con *FANCS* (*BRCA1*). Dos de estas variantes se presentaron en pacientes que desarrollaron alteraciones asociadas a malignización hematológica (c.1318G>T y c.1471-7T>G).

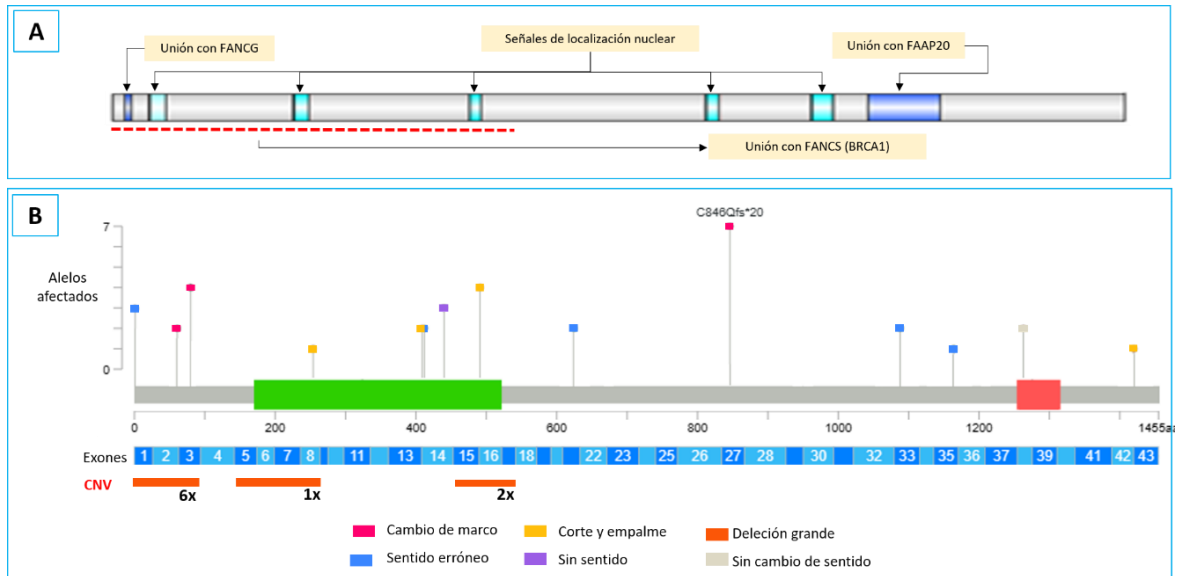


Figura 8. Heterogeneidad y distribución de las variantes descritas en los pacientes afectados en *FANCA*. A. Estructura de la proteína *FANCA*, en la que se ilustran sus señales de localización nuclear, así como las regiones de interacción con *FANCG*, *FAAP20* y *FANCS (BRCA1)*. B. Ubicación y efecto de las variantes sobre *FANCA*. Los dominios descritos en la proteína se representan como rectángulos que sobresalen del resto de su estructura (en gris); en verde: extensión del dominio N-terminal; en rojo: extensión del dominio C-terminal. En la Figura, CNV: Variante en el número de copias.

9.3 Resultados que cubren el objetivo particular 2: Caracterizar la patogenicidad de las variantes encontradas sobre sus productos proteicos mediante análisis bioinformático.

9.3.1 Análisis *in silico* del impacto de las variantes.

Una vez clasificadas las variantes, se procedió a realizar un análisis bioinformático, a fin de utilizarlo como soporte para la estimación de su patogenicidad. Para ello, en una primera etapa fueron agrupadas de acuerdo a su clase, y luego se evaluaron diversos parámetros, utilizando herramientas computacionales basadas en diferentes algoritmos para examinar su efecto. Para las variantes no sinónimas (de sentido erróneo), se valoraron tres características: propiedades físico-químicas y alineamiento múltiple de secuencias (funcionalidad), estabilidad y conservación filogenética. Como se muestra en la Tabla 7, en general, la mayoría de los programas empleados predicen un efecto negativo de los cambios sobre las proteínas evaluadas de forma consensuada, aunque también hubo variantes con estimaciones discrepantes (como p.A412V en *FANCA*). En cuanto a las variantes que conducen a cambio de marco de lectura o a codón de paro, su predicción fue más homogénea, considerándose todas como patogénicas, de acuerdo a los valores proporcionados por los programas utilizados (ver Tabla Suplementaria S2).

Tabla 7. Evaluación bioinformática de los cambios de sentido erróneo reportados en este estudio.

Gen	Cambio	Parámetro	Funcionalidad				Estabilidad		Conservación	
			SIFT 4G	Polyphen 2	MutPred2	SNAP2	MUpro	i-Mutant2	PHANTER	Consurf
FANCA	M1L	Valor	0	0.026	0.316	19	-0.57	-0.74	0.85	7
		Efecto	Deletéreo	Benigna	Benigna	Efecto	Disminuye	Disminuye	Prob. Dañina	Conservado
	A412V	Valor	0.05	0.349	0.169	-61	-0.03	-0.37	0.5	6
		Efecto	Tolerada	Benigna	Benigna	Neutro	Disminuye	Disminuye	Pos. Dañina	Promedio
	C625S	Valor	0.021	0.722	0.578	70	-1.51	-1.05	0.85	9
		Efecto	Deletéreo	Pos. Dañina	Patogénica	Efecto	Disminuye	Disminuye	Prob. Dañina	Conservado
	S1088F	Valor	0.14	0.641	0.461	92	0.67	0.63	0.5	8
		Efecto	Tolerada	Pos. Dañina	Benigna	Efecto	Incrementa	Incrementa	Pos. Dañina	Conservado
P1164S	Valor	0	0.999	0.661	50	-0.91	-1.64	0.86	9	
	Efecto	Deletéreo	Prob. Dañina	Patogénica	Efecto	Disminuye	Disminuye	Prob. Dañina	Conservado	
FANCF	M1R	Valor	0	0.999	0.901	3	-0.68	-0.24	0.89	9
		Efecto	Deletéreo	Prob. Dañina	Patogénica	Efecto	Disminuye	Incrementa	Prob. Dañina	Conservado
FANCB	H44P	Valor	0.207	0.748	0.366	30	-0.54	0.08	0.85	5
		Efecto	Tolerado	Pos. Dañina	Benigna	Efecto	Disminuye	Incrementa	Prob. Dañina	Promedio
FANCE	R371W	Valor	0.001	0.989	0.234	83	-0.82	-0.53	0.86	9
		Efecto	Deletéreo	Prob. Dañina	Benigna	Efecto	Disminuye	Disminuye	Prob. Dañina	Conservado
FANCD2	R1392P	Valor	0	0.999	0.837	95	-1.4	-1.01	0.85	9
		Efecto	Deletéreo	Prob. Dañina	Patogénica	Efecto	Disminuye	Disminuye	Prob. Dañina	Conservado

En esta tabla: Para cada programa, se muestra el valor numérico (parte superior) y el efecto del cambio (parte inferior) sobre la posición en la proteína evaluada. El rango de los valores mostrados, así como el efecto descrito, varía de acuerdo al programa utilizado (SIFT: <0.05: deletéreo; >0.05: tolerado, Polyphen 2 (modelo HumDiv): 0-0.15: benigna, 0.15-0.85: posiblemente dañina, >0.85%: probablemente dañina, MutPred2: >0.5: patogénica, <0.5: benigna, SNAP2: <-100 (neutral) hasta 100> (efecto funcional), MUpro: <0, estabilidad incrementada, >0 estabilidad disminuida, i-Mutant 2: <0, disminuye la estabilidad, >0: estabilidad aumentada, PHANTER: <0.25: benigna, >0.25-0.80: posiblemente dañino, >0.80: probablemente dañino, Consurf: 1-3: poco conservado, 4-6: intermedio (promedio), 7-9: muy conservado).

Además de categorizar las modificaciones sin sentido (de sentido erróneo) a partir de los valores estimados por cada programa, también se estudió su impacto estructural a partir de simulaciones en 3D. Por ejemplo, en el caso de la variante p.C625S sobre FANCA, la misma se produce en una posición evolutivamente muy conservada. Adicionalmente, aunque la sustitución de una cisteína por una serina sólo implica el cambio de un átomo de azufre por uno de oxígeno (Figura 9, A), modifica radicalmente las interacciones intramoleculares en esa posición, disminuyendo la estabilidad de la proteína al alterar la entropía vibracional (Figura 9, B y C).

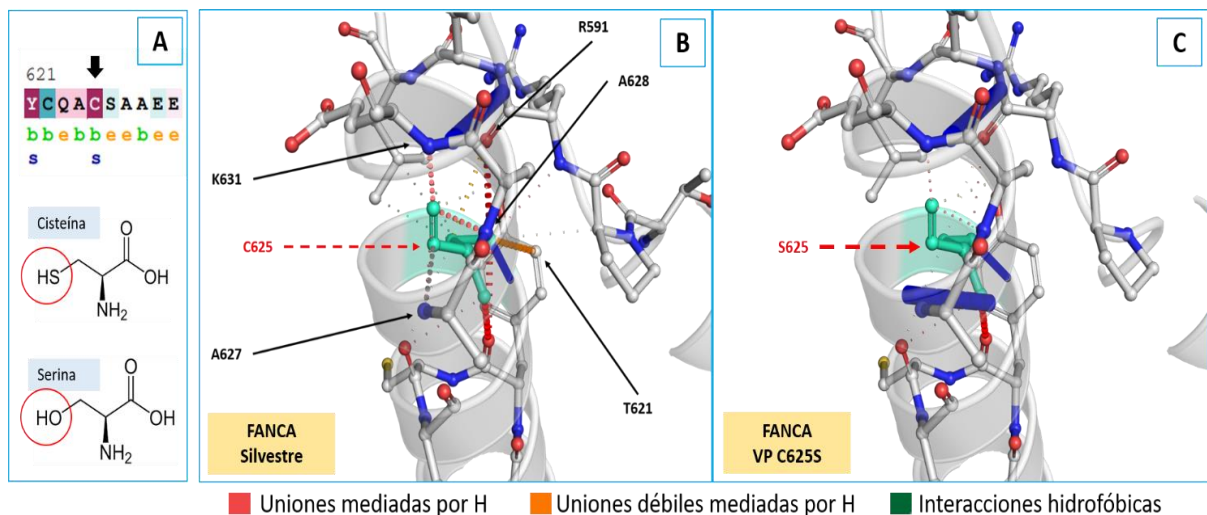


Figura 9. El cambio p.C625S altera la estabilidad de la proteína FANCA. A. En la parte superior de la figura se muestra la secuencia de aminoácidos aledaños a la cisteína ubicada en la posición 625 de FANCA (indicada con una flecha negra). La letra *b* señala que este cambio ocurre en un aminoácido orientado hacia la parte hidrofóbica de la proteína, mientras que la letra *s* hace referencia a su carácter estructural. En la parte inferior, se muestra la estructura del aminoácido silvestre (cisteína, C) y del modificado (serina, S). B. Interacciones intramoleculares (mostradas como puntos de color discontinuos) y localización espacial de los aminoácidos cercanos a la cisteína (en color verde), en una hélice α de FANCA. En azul, átomos de N, en rojo, átomos de Oxígeno, en blanco: átomos de Carbono. C. Misma representación que en B, cuando se reemplaza la cisteína por serina (en verde), en donde se aprecia la pérdida de interacciones en la posición evaluada.

Finalmente, en el caso de las variantes de corte y empalme, la mayoría de ellas (cuatro de cinco, 80%) se localizó en sitios canónicos (aceptor y receptor), cuya conservación es crítica para el mecanismo de corte y empalme. En estos casos los programas empleados predijeron la perturbación de estas posiciones en las isoformas primarias, ocasionando saltos de exones (*skipping*) y sitios crípticos (ver Tabla suplementaria S3). De manera interesante, también se encontró una variante que probablemente afecta otro sitio de regulación de este proceso (FANCA: c.1471-7T>G, localizada en el intrón 15), ya que se ubica en una secuencia considerada como punto de ramificación ("*branch point*"), cuya localización y estructura de motivos puede diferir dependiendo del transcrito. Para analizar su impacto, se utilizaron dos programas (HSF y RNABPS). La información utilizada para caracterizarla se ilustra en la Figura 10. En ella, podemos observar que el cambio ocasionado por la variante, para el punto de ramificación predicho, presenta una alta probabilidad de perturbar dicha posición. Es importante mencionar que uno de los individuos afectado por esta variante desarrollo SMD.

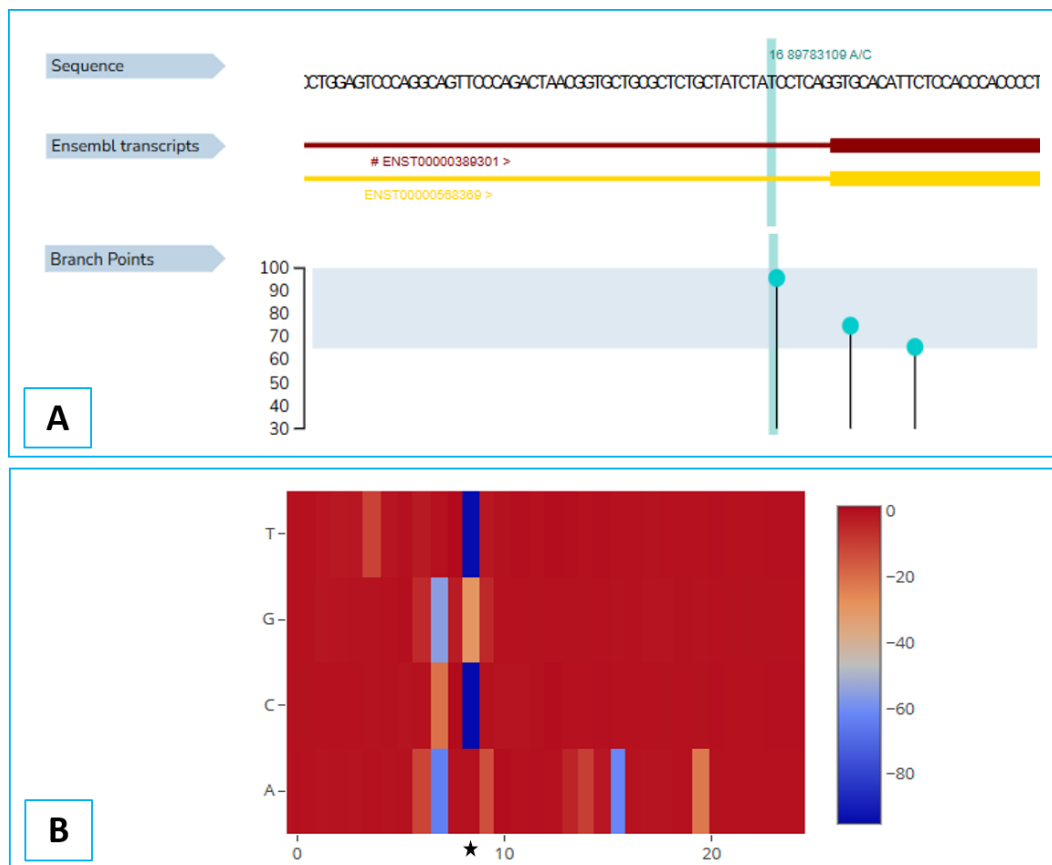


Figura 10. Impacto del cambio *FANCA*: c.1471-7T>G sobre un punto de ramificación en el intrón 15, asociado al proceso de corte y empalme. A. Secuencia genómica y posición de la variante (denotada con una sombra verde), y su ubicación correspondiente en los transcritos de *FANCA*. En la parte inferior se observa la predicción de los puntos de ramificación. La altura hace referencia a la probabilidad para cada uno de los puntos de ramificación predichos sobre el transcrito primario. B. Impacto del cambio de nucleótido sobre el punto de ramificación. La estrella hace referencia a la posición evaluada, mientras que el mapa de calor, a la derecha de la figura, indica el efecto de los cambios de nucleótido para el punto de ramificación predicho (en este caso, A por C).

Debido a que el estudio del efecto de una variante a través de herramientas bioinformáticas forma parte de los criterios sugeridos por la ACMG-AMP para su la clasificación, estos resultados se integraron a otras evidencias de patogenicidad (como la frecuencia alélica, prevalencia en individuos afectados, evidencia de estudios funcionales, patrón de cosegregación en las familias estudiadas, entre otras) para establecer un veredicto. Los resultados de este ejercicio se resumen en la Tabla 8, y se sintetizan en la Figura 11, en donde se observa la elevada proporción de variantes clasificadas como patogénicas o probablemente patogénicas, que en conjunto abarcan >75%. Sin embargo, resalta la ausencia de estos tipos de variante en 2 genes (*FANCE* y *FANCD2*), en los que se presentó recurrentemente variantes de significado incierto. El uso de herramientas bioinformáticas para la caracterización de las variantes de efecto fundador se contextualiza y describe detalladamente en la siguiente sección.

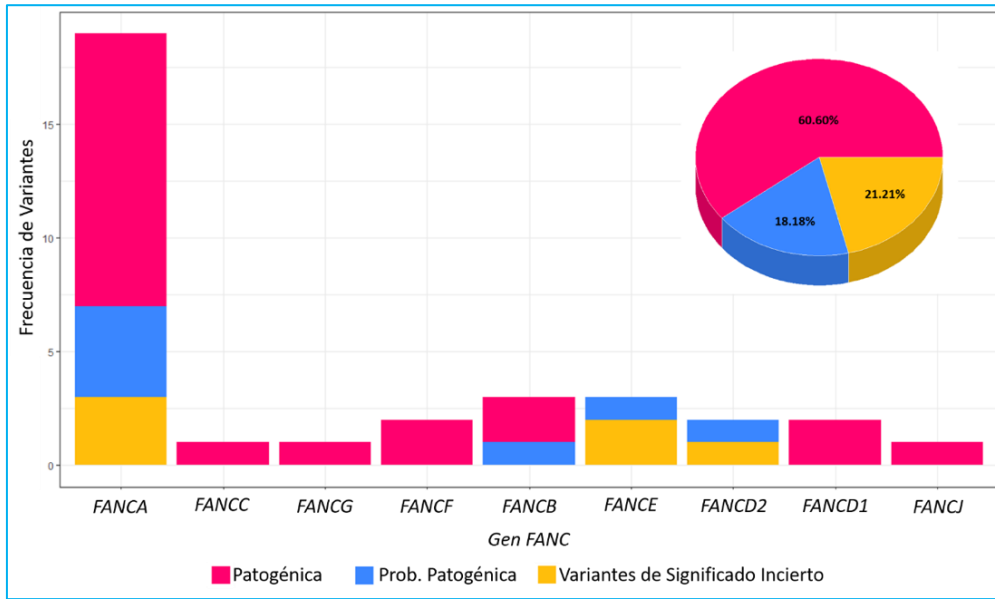


Figura 11. Distribución de las variantes encontradas en este estudio, de acuerdo a su grado de patogenicidad, luego de aplicar los criterios de la ACMG. En la parte superior derecha de la figura se muestra la frecuencia de las variantes, con respecto a la totalidad descrita en este estudio.

Tabla 8. Clasificación de variantes de acuerdo a su patogenicidad, utilizando los criterios sugeridos por la Colegio Americano de Genética Médica (ACMG, por sus siglas en inglés).

Variante	Tipo	Efecto en la proteína	Alelos Afectados (pacientes)	Criterios ACMG	Clasificación ACMG
FANCA					
c.(?-42)_(283+1_284-1)del	DG	p.?	6x (3)	PVS1,PM2,PP5	Patogénica
c.1235C>T	SE	p.Ala412Val	2x (2)	BP4,BP1,PS2,PM6	Significado I
c.3263C>T	SE	p.Ser1088Phe	2x (2)	BP1,BP4,PP5,PS3	Significado I
c.2535_2536delCT	CM	p.Cys846GlnfsTer20	7x (4)	PVS1,PM2,PP5	Patogénica
c.(522+1_523-1)_(792+1_793-1)del	DG	p.?	1x (1)	PVS1,PM2,PP5	Patogénica
c.1226-2A>C	CyE	NA	2x (2)	PVS1,PM2,PP5	Patogénica
c.(1359+1_13601)_(1626+1_1627-1)del	DG	p.?	2x (2)	PVS1,PM2,PP5	Patogénica
c.1874G>C	SE	p.Cys625Ser	2x (2)	PM2,PP3, PP5	Significado I
c.182_184delTGCinsGG	CM	p.Leu61TrpfsTer4	2x (2)	PVS1,PM2	P. Patogénica
c.238del	CM	p.Cys80ValfsTer15	4x (2)	PVS1,PM2,PP5	Patogénica
c.1A>C	SE	p.Met1Leu	3x (2)	PVS1,PM2,PP5	Patogénica
c.1318G>T	SS	p.Glu440Ter	3x (2)	PVS1,PM2,PP5	Patogénica
c.4260+1G>A	CE	NA	1x (2)	PVS1,PM2,PP5	Patogénica
c.827-1G>A	CE	NA	1x (2)	PVS1,PM2	P. patogénica
c.3490C>T	SE	p.Pro1164Ser	1X (1)	PM2,PM5,PP3,PP5	Significado I
c.3788_3790del	SCS	p.Phe1263del	2x (1)	PM1,PM2,PP3,PP5	Patogénica
c.1471-7T>G	CE	NA	4x (1)	PM2,PP3	P.patogénica
FANCC					
c.67delG	CM	p.Asp23IlefsTer23	6x (3)	PVS1,PM2,PP5	Patogénica
FANCG					
c.511-3_511-2delCA	CE	NA	6x (3)	PVS1,PM2,PP3,PP5	Patogénica
FANCF					
c.2T>G	SE	p.Met1Arg	4x (2)	PVS1,PM2,PP3,PP5	Patogénica
c.84del	CM	p.Ala29ProfsTer52	2x (1)	PVS1,PM2,PP5	Patogénica
FANCB					
c.354delT	CM	p.Phe118LeufsTer5	2x (2)	PVS1,PM2	Patogénica
c.131A>C	SE	p.His44Pro	1x (1)	PM2,BP4	Significado I
c.1437G>A	SS	p.Trp479Ter	1x (1)	PVS1,PM2,PP5	Patogénica

En esta tabla: respecto al tipo de variante: CE: corte y empalme, CM: cambio de marco de lectura, DG: delección grande, SCS: sin cambio de sentido, SE: sentido erróneo, SS: sin sentido. Respecto a la proteína: p.?: Ausencia en la producción de la proteína. NA: no aplica (es decir, es un cambio que no involucra la secuencia codificante). Criterios de la ACMG: BP: criterios de soporte para variante benigna, PM: criterios para asignar impacto patogénico moderado a una variante, PP: criterios de soporte para efecto patogénico, PVS1: variante que conduce a la producción nula de proteína o causa pérdida de su función. La numeración en cada categoría no hace referencia a la "gravedad" de un cambio; se indica como una referencia que ayuda a definir diferentes características que se evalúan durante la asignación de patogenicidad (para información adicional, consultar Richards et al, 2015). Las variantes escritas en rojo fueron encontradas en pacientes que desarrollaron neoplasias.

Tabla 8 (Continuación). Clasificación de variantes de acuerdo a su patogenicidad, utilizando los criterios sugeridos por la Colegio Americano de Genética Médica (ACMG, por sus siglas en inglés).

Variante	Tipo	Efecto en la proteína	Alelos Afectados (pacientes)	Criterios ACMG	Clasificación ACMG
<i>FANCE</i>					
c.-12_2del	DI	p.?	1x (1)	PVS1,PM2	Patogénica
c.855+178del	DI	p.?	1x (1)	PM2,BP4	Significado I
c.1111C>T	SE	p.Arg371Trp	2x (1)	PVS1,PM2,PP5	Patogénica
<i>FANCD2</i>					
c.3156_3157insGGAC	CM	p.Gln1053GlyfsTer31	1x (1)	PVS1,PM2	P. patogénica
c.4175G>C	SE	p.Arg1392Pro	1x (1)	PM2,PP3	Significado I
<i>FANCD1</i>					
c.658_659del	CM	p.Val220IlefsTer4	1x (1)	PVS1,PM2,PP5	Patogénica
c.3264dup	CM	p.Gln1089SerfsTer10	1x (1)	PVS1,PM2,PP5	Patogénica
<i>FANCI</i>					
c.2392C>T	SS	p.Arg798Ter	2x (1)	PVS1,PP5,PM2	Patogénica

En esta tabla: respecto al tipo de variante: CE: corte y empalme, CM: cambio de marco de lectura, DG: deleción grande, SCS: sin cambio de sentido, SE: sentido erróneo, SS: sin sentido. Respecto a la proteína: p.? : Ausencia en la producción de la proteína. NA: no aplica (es decir, es un cambio que no involucra la secuencia codificante). Criterios de la ACMG: BP: criterios de soporte para variante benigna, PM: criterios para asignar impacto patogénico moderado a una variante, PP: criterios de soporte para efecto patogénico, PVS1: variante que conduce a la producción nula de proteína o causa pérdida de su función. La numeración en cada categoría no hace referencia a la “gravedad” de un cambio; se indica como una referencia que ayuda a definir diferentes características que se evalúan durante la asignación de patogenicidad (para información adicional, consultar Richards et al, 2015). Las variantes escritas en rojo fueron encontradas en pacientes que desarrollaron neoplasias.

9.3.2 Caracterización de variantes de efecto fundador.

9.3.2.1 *FANCG*: c.511-3_511-2delCA

Esta variante, encontrada en tres pacientes (etiquetados como FANC32, FANC143 y FANC155), no había sido reportada previamente en pacientes con AF. Esta delección de dos bases, detectada a través de un panel de SNG, y corroborada posteriormente por secuenciación de Sanger (Figura 12, A y B) se ubica entre el borde del intrón cuatro y el inicio del exón cinco de *FANCG*, e involucra cambios en el sitio canónico aceptor, necesario durante el proceso de corte y empalme.

Luego de caracterizar el impacto de esta alteración a través de herramientas bioinformáticas, se estableció en conjunto una predicción *in silico* que indicaba la perturbación de este sitio aceptor y promueve la pérdida (“*skipping*”) del exón cinco, que posee 131 pares de bases. Esto a su vez tiene como consecuencia la generación de un corrimiento del marco de lectura, que conduce a un codón de paro prematuro. La proteína trunca producida (p.S171Vfs5*), se encuentra afectada desde la primera región de tetratricopéptidos de *FANCG* (Figura 12, C). Adicionalmente, la frecuencia de este cambio (rs1491369358) es muy baja (1/251,480 alelos), siendo este un criterio de apoyo distinto, que sugiere la patogenicidad de la variante.

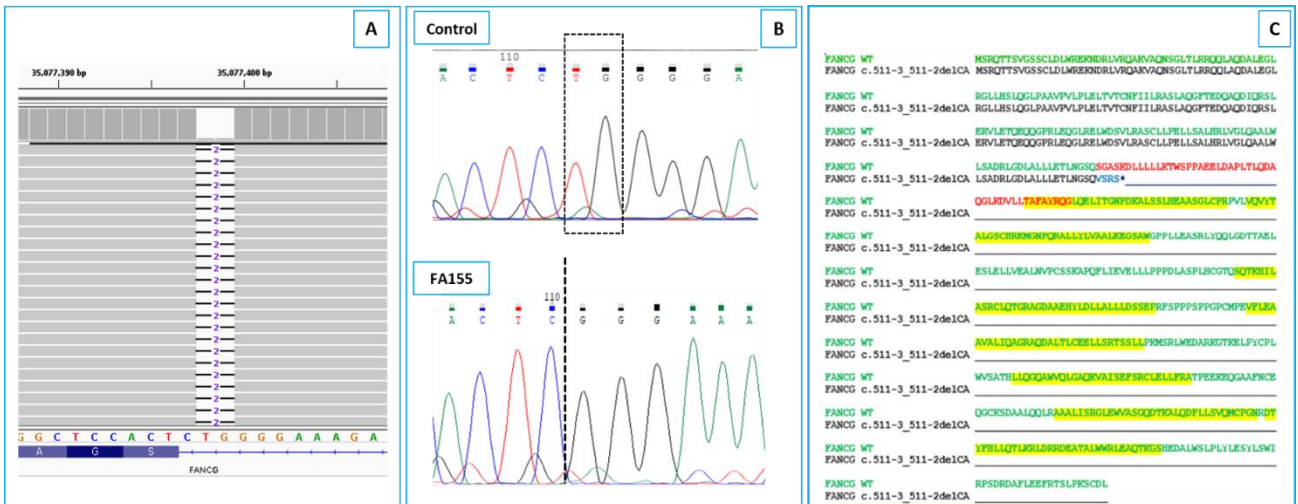


Figura 12. El análisis del efecto de la variante *FANCG*: c.511-3_511-2delCA predice la inducción de la pérdida del exón 5. A. Imagen de la delección de dos bases encontrada en los tres pacientes (archivo .bam). B. Corroboración de la variante por secuenciación de Sanger; la línea punteada en el electroferograma inferior señala la pérdida de las bases T y G en uno de los afectados (FA155). C. Efecto de la remoción del exón 5 sobre la proteína *FANCG*. En rojo, aminoácidos codificados por el exón 5, en amarillo: aminoácidos que pertenecen a las regiones de tetratricopéptidos (TPRs).

Durante la revisión de los expedientes de los tres pacientes con esta alteración, nos percatamos que provenían de la misma localidad, ubicada en una zona remota de la Sierra Norte, en el Estado de Oaxaca. Los mismos se identificaban como no relacionados (aunque compartían un apellido), y pertenecían al grupo indígena *Mixe*. Debido a estos antecedentes, se decidió investigar la posibilidad de que se tratara de una VP fundadora y que compartieran bloques de haplotipos. Para validar esta presunción, las muestras de DNA se estudiaron con microarreglos de alta densidad.

A partir de su análisis se encontraron dos LSCH compartidos entre los pacientes, encontrándose una de ellas en una región del brazo corto del cromosoma nueve, con un tamaño de 9018 Kb que abarcaba el *locus* donde se encuentra *FANCG*, además de otro bloque en el brazo largo del cromosoma cinco. La exploración de estas regiones nos permitió definir una posible relación entre los pacientes, que se evaluó a través de la estimación del grado de parentesco (valor F, ver Tabla 9), cuyo resultado refleja una posible relación de tercer y cuarto grado.

Tabla 9. Análisis de parentesco entre los tres pacientes con genotipo *FANCG*.

Código del paciente	Extensión de los fragmentos contiguos (LCSH > 3 Mb) (kb)	Porcentaje de componente nativo	Homocigosidad en autosomas (%)	F	Probable relación parental
FANC32	140214361	79.17%	4.9%	0.0487	Cuarto grado
FANC143	203913825	85.64%	7.1%	0.0708	Tercer grado
FANC155	249262167	88.53%	8.7%	0.0865	Tercer grado

Para la búsqueda del haplotipo alrededor de *FANCG*, se utilizó la LSCH compartida por los tres pacientes en el cromosoma nueve. En esta zona se detectaron 544 marcadores de nucleótido único, de los cuales 439 (79.41%) presentaban la misma configuración alélica. Cuando se buscaron las frecuencias alélicas de estos marcadores a nivel mundial, así como las reportadas para población mexicana en el Proyecto de los 1000 Genomas (1KGP), encontramos que 99.32% de los marcadores eran coincidentes. A partir de esta información se pudo construir un haplotipo, formado por 30 marcadores alrededor de *FANCG*. La presencia de este haplotipo respaldó el efecto fundador de la variante.

Finalmente, para apoyar que el efecto fundador se originó en México, se realizó un análisis de ancestría global. Para ello se utilizaron todos los marcadores compartidos por los tres pacientes *Mixes*. Para considerar el aporte de diferentes poblaciones en el pasado sobre la gran diversidad de la población mexicana, y reflejar su alto grado de mestizaje, se seleccionaron como poblaciones parentales del 1KGP a los Nativos Americanos (NAT), Residentes de Utah con ancestría de Europa del Norte y Occidental; y a la población Yoruba (YRI), procedente de Ibadan, Nigeria.

Los resultados indican que la población mexicana presentaba un patrón de ascendencia con predominancia de los componentes nativo-americanos y europeo (NAT: 46.59% y CEU: 48.44%, respectivamente) (ver Figura 13, A). No obstante, el aporte del componente NAT en los pacientes de origen *Mixe* fue mucho mayor (79%-88%). A pesar de este contraste, la diversidad mostrada en la población mexicana se conserva en los pacientes estudiados.

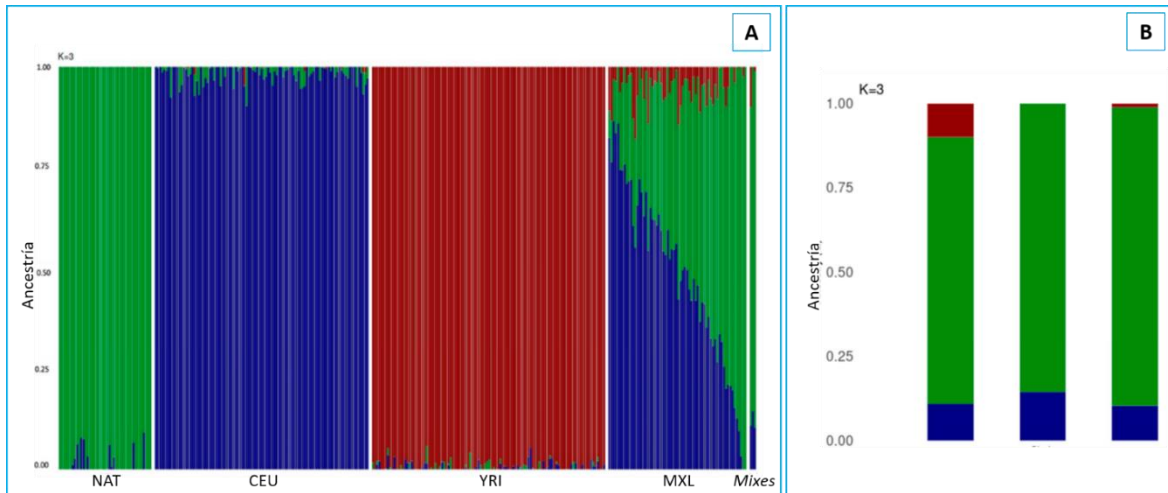


Figura 13. Los pacientes *Mixes* con Anemia de Fanconi presentan un patrón de ascendencia que refleja las características de la diversidad genética de la población mexicana. A. Diagrama de barras que muestra la inferencia de la ascendencia global para los individuos *Mixes* y mexicanos utilizando ADMIXTURE. Los paneles de referencia poblacional en este caso incluyen NAT (nativos americanos), CEU (europeos del norte) y YRI (yorubas) (modelo K = 3). Cada individuo se representa como una barra vertical. Los colores muestran el porcentaje de ascendencia asignado a cada grupo para cada individuo. Los colores verde, azul y rojo indican las poblaciones NAT, CEU y YRI, respectivamente. Los dos paneles situados en el extremo derecho de la figura muestran los componentes de ascendencia tripartita para la población mexicana mezclada (MXL) y los pacientes con AF de origen *Mixe* (últimas tres barras). (B) Gráfico de barras detallado de los tres pacientes (de izquierda a derecha: FANC32, FANC143 y FANC155).

9.3.2.2. FANCC: c.67delG.

La variante descrita en *FANCG* no fue la única con efecto fundador encontrada en nuestro grupo de pacientes con AF, ya que también se detectó otra en tres pacientes provenientes del estado de Tamaulipas (etiquetados como FANC152, FANC153 y FANC154). El caso índice, hijo de padres consanguíneos, procedía de una comunidad de ascendencia Menonita. Al extender la prueba de diagnóstico citogenético a sus 10 hermanos, dos de ellos resultaron afectados con AF (Figura 14, A). En este grupo étnico, surgido en los Países Bajos y que ha tenido un patrón migratorio que le ha permitido asentarse en otros países, se ha reportado previamente una VP predominante para AF (*FANCC: c.67delG*; p. Asp231IlefsTer23). Con este contexto, se decidió buscar dicha

alteración de forma dirigida en esta familia, a través de secuenciación de Sanger. Se encontró la VP previamente reportada, en los tres individuos afectados y adicionalmente se encontraron seis portadores (Figura 14, A y B).

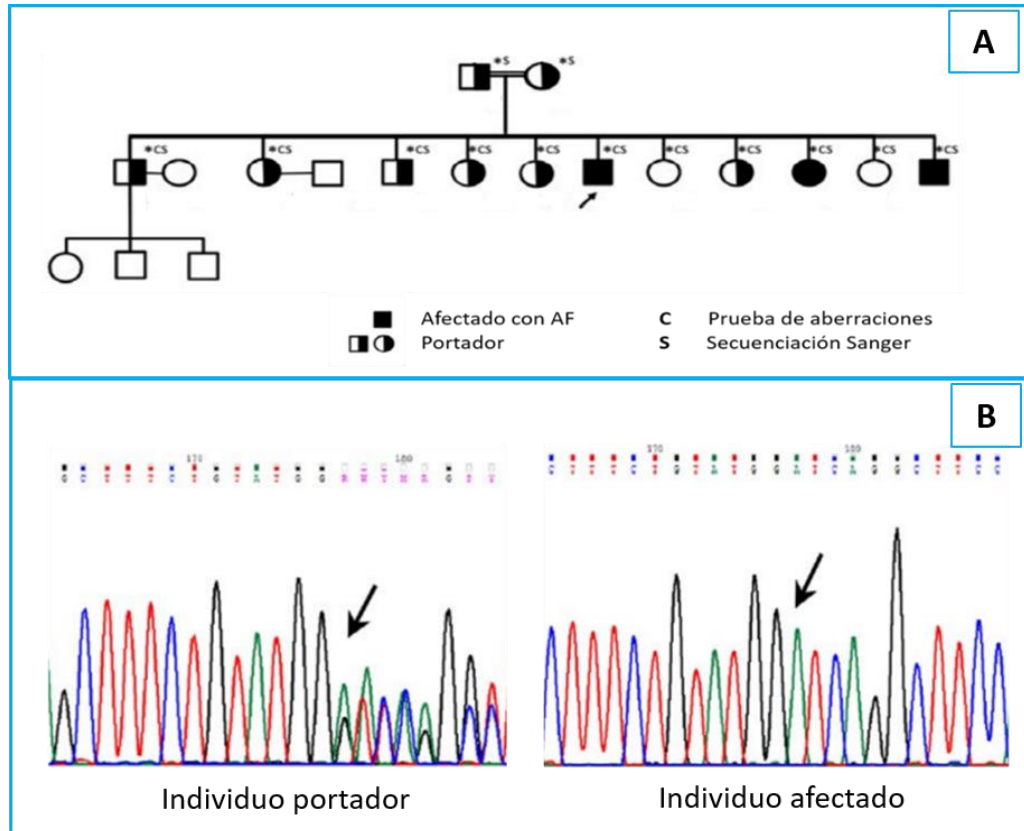


Figura 14. Validación de la variante *FANCC*: c.67delG por secuenciación de Sanger dirigida. A. Árbol genealógico de la familia Menonita estudiada. En la figura, los cuadrados representan pacientes de sexo masculino, mientras que los círculos corresponden a pacientes de sexo femenino. B. Electroferogramas de la secuenciación de Sanger, la flecha negra señala la posición de la deleción.

Para corroborar el efecto fundador de la variante se utilizaron dos SNVs aledañas a la ubicación de la variante (rs2277182 y rs016013, presentes en la base de datos Refseq), que habían sido empleadas en reportes anteriores para el análisis de haplotipos en individuos Menonitas ubicados en Manitoba, Canadá (de Vries et al, 2012, ver Figura 15). Este estudio confirmó la homociguidad de los pacientes con AF y de sus padres para los dos marcadores empleados (Tabla 10), evidenciando que tanto el alelo alterado como el silvestre segregaban de manera conjunta con estos SNVs.

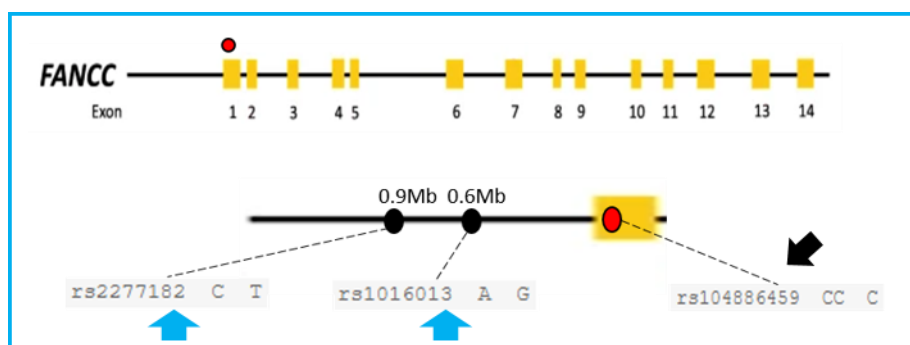


Figura 15. Uso de SNVs para la determinación de haplotipo en *FANCC*. En la parte superior de la figura, cada recuadro amarillo representa un exón de *FANCC*. La localización de la variante se representa con un círculo rojo. En la parte inferior, las flechas azules señalan los SNVs utilizados para la búsqueda del haplotipo, mientras que la flecha negra señala a la VP evaluada (*FANCC*:c.67delG)

Tabla 10. Configuración alélica de los marcadores para el análisis de haplotipos en *FANCC*.

Individuos	rs2277182	rs016013	rs104886459
	C/T	A/G	delG
Mamá	C/C	G/G	-/CC
Papá	C/C	G/G	-/CC
FANC152	C/C	G/G	delG/delG
FANC153	C/C	G/G	delG/delG
FANC154	C/C	G/G	delG/delG

9.4. Resultados que cubren el objetivo 3. Correlacionar el tipo y localización de las variantes patogénicas encontradas con las características clínicas de los pacientes.

9.4.1. Correlación entre el genotipo y el fenotipo físico.

De los 41 pacientes con genotipo, en 37 casos se obtuvieron datos fenotípicos a partir de la exploración física (Figura 16) y de la revisión de expedientes. Para visualizar las asociaciones entre la frecuencia de estos rasgos y el genotipo, se generaron mapas de calor para cada característica evaluada. Respecto a la asociación VACTERL-H, las alteraciones en las extremidades fueron las más frecuentes, encontrándose en 56% de los pacientes, seguidas de las renales (26.8%) (Figura 17, A). No se reportaron datos indicativos de fístula traqueo-esofágica o atresia esofágica en este grupo de estudio. Aunque sólo dos pacientes presentaron tres o más de los rasgos necesarios para presentar esta asociación (4.87%), en 80.48% de los casos se observó al menos una de sus anomalías.

En relación al acrónimo PHENOS, las alteraciones en la pigmentación de la piel fueron las más recurrentes (65.85%, Figura 16, A y Figura 17, B), siendo este el rasgo fenotípico reportado con mayor frecuencia entre todas las características físicas evaluadas. Le siguieron en frecuencia para este acrónimo las vinculadas a estatura baja (44.4%, Figura 16, B y Figura 17, B). En ocho (19.51%) pacientes se presentaron 4 o más de los rasgos descritos para este acrónimo. No se reportaron pacientes que tuvieran de manera combinada VACTERL-H y PHENOS.

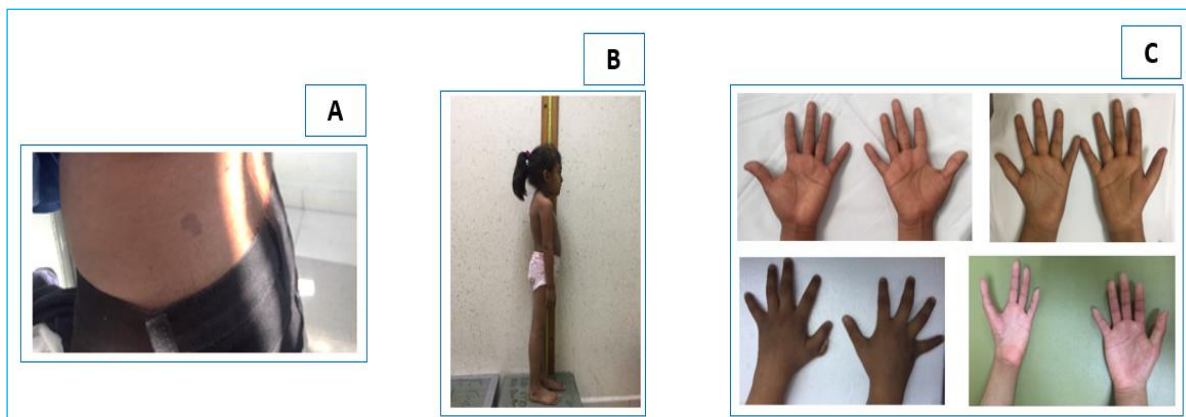


Figura 16. Características físicas reportadas con mayor frecuencia en los pacientes con Anemia de Fanconi con asignación de genotipo. A. Alteraciones en la pigmentación de la piel, en la que se observa una mancha *café au lait*. B. Talla baja. C. Alteraciones en las extremidades superiores, en las que se observan hipoplasia tenar y de dedos, ausencia de pulgares y polidactilia preaxial.

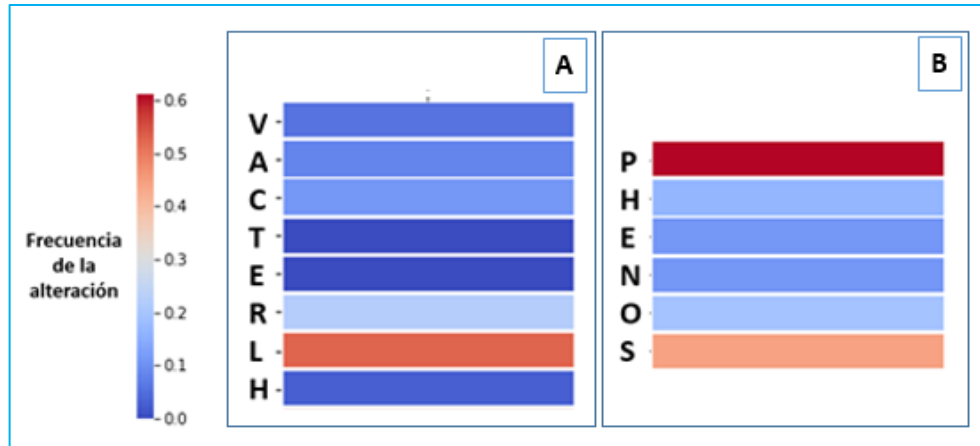


Figura 17. Frecuencia de las características descritas para la asociación VACTERL-H y el acrónimo PHENOS, en pacientes con Anemia de Fanconi y genotipo. A. Proporción de las alteraciones en la asociación VACTERL-H. B. Proporción de las alteraciones en el acrónimo PHENOS. En la figura, las letras en el eje vertical se organizan de acuerdo a la abreviatura de las alteraciones descritas para el acrónimo estudiado: V (vertebrales), A (anales), C (cardíacas), T (fístula tráqueo-esofágica), E (atresia esofágica), R (renales), L (del inglés limbs, anomalías en extremidades), H (hidrocefalia); P (pigmentación), H (del inglés small head, cabeza pequeña), E (del inglés small eyes, ojos pequeños), N (anomalías en el sistema nervioso central), O (alteraciones otológicas) y S (del inglés short stature, talla baja). La frecuencia de cada rasgo se representa en un “mapa de calor”, en el que el color rojo cálido corresponde a lo encontrado en mayor proporción, en contraste con el color azul rey, asignado a los rasgos menos habituales.

Como primera aproximación en la búsqueda de la correlación genotipo-fenotipo, se exploró la asociación entre el tipo de variante y las alteraciones descritas en VACTERL-H y PHENOS. Las variantes se clasificaron de acuerdo a su impacto sobre las proteínas de la vía FA/BRCA, sin considerar el gen *FANCA* afectado, en dos grupos: hipomórficas (conservan actividad residual) y nulas (producción comprometida o ausencia total). Los resultados de esta clasificación muestran la predominancia en ambos grupos de alteraciones en las extremidades para la asociación VACTERL-H, aunque sin diferencias significativas. Para el acrónimo PHENOS, las alteraciones en la estatura y en la pigmentación de la piel fueron las más frecuentes. No obstante, para esta última característica si hubo una diferencia significativa entre ambos grupos (Prueba Exacta de Fisher con $\alpha=0.05$, p-valor=0.016).

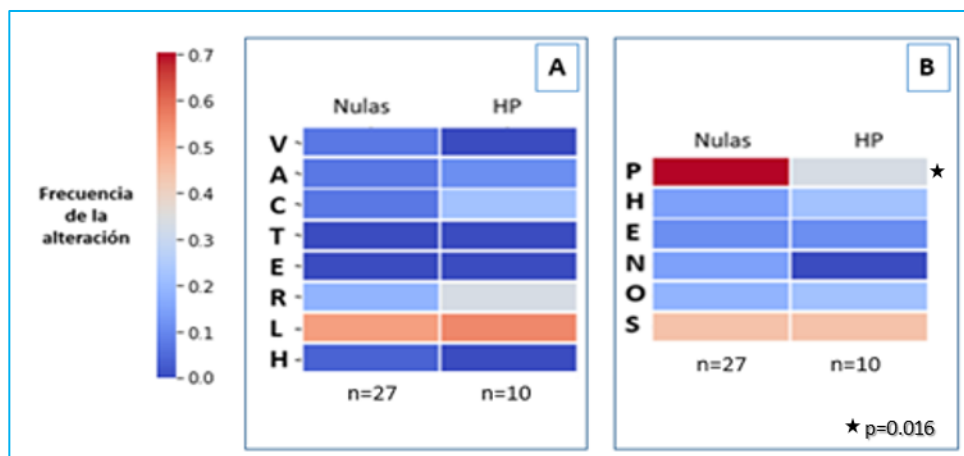


Figura 18. Correlación entre el tipo de variante encontrada (nulas vs hipomórficas (HP)) y el fenotipo físico. A. Proporción de las alteraciones en la asociación VACTERL-H. B. Proporción de las alteraciones en el acrónimo PHENOS. El asterisco señala diferencias significativas entre ambos grupos, para la característica evaluada. n=número de pacientes por grupo evaluado. En la figura, las letras en el eje vertical se organizan de acuerdo a la abreviatura de las alteraciones descritas para el acrónimo estudiado: V (vertebrales), A (anales), C (cardíacas), T (fístula tráqueo-esofágica), E (atresia esofágica), R (renales), L (del inglés *limbs*, anomalías en extremidades), H (hidrocefalia); P (pigmentación), H (del inglés *small head*, cabeza pequeña), E (del inglés *small eyes*, ojos pequeños), N (anomalías en el sistema nervioso central), O (alteraciones otológicas) y S (del inglés *short stature*, talla baja). La frecuencia de cada rasgo se representa en un “mapa de calor”, en el que el color rojo cálido corresponde a lo encontrado en mayor proporción, en contraste con el color azul rey, asignado a los rasgos menos habituales.

Por otra parte, también se estudió la asociación entre las anomalías físicas y la posición del gen afectado de acuerdo a la descripción clásica de la vía FA-BRCA, formada por tres complejos (“río arriba”, “FANCI-FANCD2” y “río abajo”). Sin embargo, debido a que los pacientes afectados en genes que codifican para proteínas integradas en el complejo “río arriba” representan 92.68% del total (38 de 41), decidimos examinar si la división de este ensamblaje multiproteico en tres subcomplejos (AG20, BL100 y CEF), propuesta a nivel molecular, representaba alguna asociación a nivel del fenotipo físico. Se obtuvieron datos fenotípicos derivados de los expedientes en 36 de estos casos. La Figura 19 muestra esta comparación. A pesar de que visualmente en el grupo BL100 se observan diferencias para PHENOS y para algunas características de VACTERL-H, en relación a los otros dos subcomplejos, este resultado debe tomarse con precaución, ya que el número de pacientes es muy pequeño (el grupo sólo está representado por dos pacientes con VP en *FANCB*). Asimismo, a pesar de que la proporción de individuos afectados en extremidades tiende a ser mayor en el grupo CEF con respecto al grupo AG20, esta diferencia no resultó significativa (Prueba exacta de Fisher con $\alpha=0.05$, $p=0.680$).

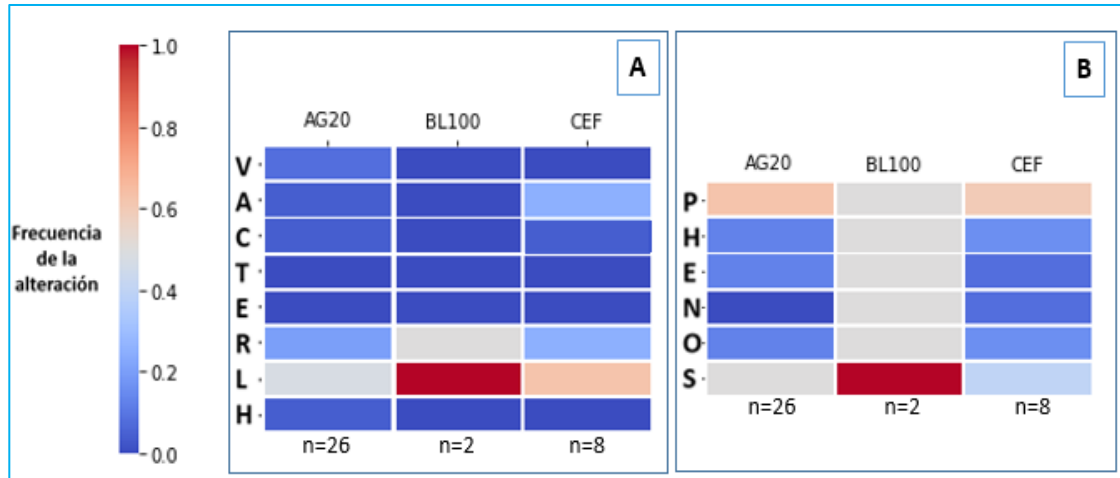


Figura 19. Correlación entre los subcomplejos del complejo central (AG20, BL100 y CEF) y el fenotipo físico. A. Proporción de alteraciones en la asociación VACTERL-H. B. Proporción de alteraciones en el acrónimo PHENOS. n=número de pacientes por grupo evaluado. En la figura, las letras en el eje vertical se organizan de acuerdo a la abreviatura de las alteraciones descritas para el acrónimo estudiado: V (vertebrales), A (anales), C (cardíacas), T (fístula tráqueo-esofágica), E (atresia esofágica), R (renales), L (del inglés limbs, anomalías en extremidades), H (hidrocefalia); P (pigmentación), H (del inglés small head, cabeza pequeña), E (del inglés small eyes, ojos pequeños), N (anomalías en el sistema nervioso central), O (alteraciones otológicas) y S (del inglés short stature, talla baja). La frecuencia de cada rasgo se representa en un “mapa de calor”, en el que el color rojo cálido corresponde a lo encontrado en mayor proporción, en contraste con el color azul rey, asignado a los rasgos menos habituales.

Por otra parte, en este trabajo se observó que el gen alterado con mayor frecuencia en nuestra cohorte fue *FANCA* (ver Figura 5), abarcando más de la mitad de los casos, mientras que la distribución de afectados en otros genotipos es relativamente homogénea, siendo *FANCB* el segundo más afectado en los pacientes (cerca al 10%). En adición, el espectro de variantes en *FANCA* fue más heterogéneo, encontrándose prácticamente todos los tipos descritos, en contraste con lo reportado en los otros genes *FANC*. Por estas razones, nos propusimos analizar si una separación de grupos de acuerdo a las frecuencias (sugiriendo, entonces, pacientes con variantes en *FANCA* contra aquellos con variantes en otro gen *FANC*) también podría señalar alguna tendencia relevante al estudiar las manifestaciones físicas. Esta comparación se observa en la Figura 20.

La descripción de acuerdo al mapa de calor sugiere que ambos grupos presentan proporciones de frecuencias similares en la mayoría de las características evaluadas. Se observa una tendencia mayor de alteraciones renales y de extremidades en los pacientes con genotipo *FANCA*, sin embargo, estas diferencias no fueron significativas (Prueba exacta de Fisher con $\alpha=0.05$, $p=0.143$ y $p=0.73$, respectivamente).

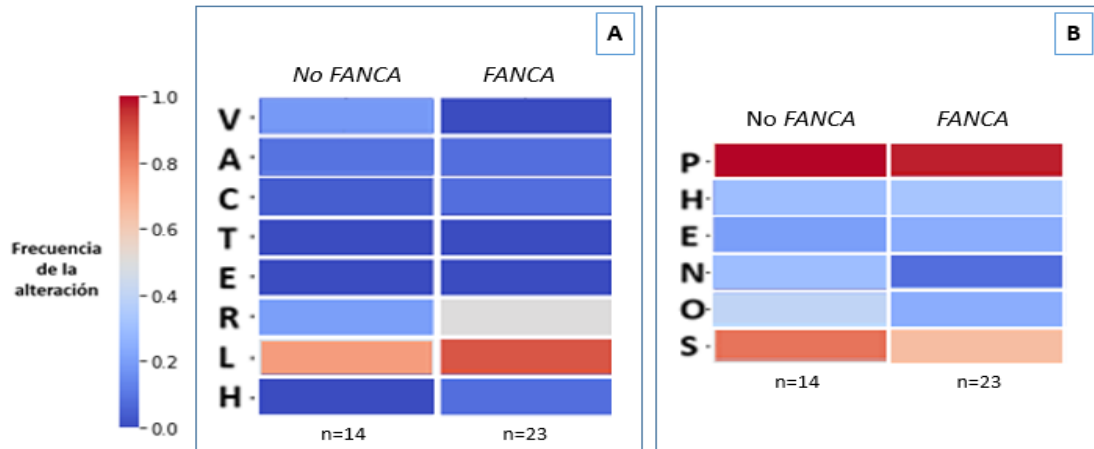


Figura 20. Correlación entre pacientes con genotipo FANCA contrastada con aquellos con genotipo diferente a este gen (grupo No FANCA), en cuanto al fenotipo físico. A. Proporción de las alteraciones en la asociación VACTERL-H. B. Proporción de las alteraciones para el acrónimo PHENOS. n=número de pacientes por grupo evaluado. En la figura, las letras en el eje vertical se organizan de acuerdo a la abreviatura de las alteraciones descritas para el acrónimo estudiado: V (vertebrales), A (anales), C (cardíacas), T (fístula tráqueo-esofágica), E (atresia esofágica), R (renales), L (del inglés limbs, anomalías en extremidades), H (hidrocefalia); P (pigmentación), H (del inglés small head, cabeza pequeña), E (del inglés small eyes, ojos pequeños), N (anomalías en el sistema nervioso central), O (alteraciones otológicas) y S (del inglés short stature, talla baja). La frecuencia de cada rasgo se representa en un “mapa de calor”, en el que el color rojo cálido corresponde a lo encontrado en mayor proporción, en contraste con el color azul rey, asignado a los rasgos menos habituales.

9.4.2. Correlación entre el genotipo y el fenotipo hematológico de los pacientes.

Además de la revisión del fenotipo físico de los pacientes, también se analizaron datos de exámenes sanguíneos, a fin de obtener información sobre la evolución y seguimiento de su condición hematológica. La falla medular (FM), cuando se presentó, fue clasificada en tres niveles: leve, moderada o severa. Se obtuvieron datos hematológicos de las fechas más cercanas al momento del diagnóstico citogenético de AF en 34 (82.92%) casos; la mitad de ellos presentó genotipo FANCA. Se observó una distribución similar de los pacientes afectados de acuerdo al nivel de FM; y en total estos duplicaron los casos sin falla medular (Tabla 11). Asimismo, en los pacientes con FM, la forma severa se presentó más temprano que las formas leves y moderadas, que presentaron un promedio de edad semejante.

Tabla 11. Clasificación de la falla medular en 34 pacientes con Anemia de Fanconi y genotipo determinado.

Condición de la falla medular	Número de pacientes	Edad al diagnóstico (promedio, en años)	Rango de edades
Sin falla medular	12	6.09	0.29-15.3
Leve	8	7.63	1.4-11.7
Moderada	6	7.64	5.17-17.07
Severa	8	6.75	0.36-10.40

La información derivada de la genotipificación se utilizó para conocer si existían potenciales correlaciones con el fenotipo de FM. Inicialmente se evaluó si el tipo de variante (nulas vs hipomórficas) estaba asociado a la progresión de la FM. Aunque la proporción de variantes nulas predomina sobre las hipomórficas, la distribución de ambos tipos fue homogénea entre los grupos estudiados (Figura 21, A); por lo que el grado de FM no tiene relación con el tipo de variante.

Por otra parte, en relación al gen alterado, el grupo de pacientes sin FM, así como aquellos con FM severa, presentaron una distribución heterogénea entre sus genotipos, en contraste con las formas leves y moderadas, en las que predominó *FANCA*. Sin embargo, destaca que sólo uno de los pacientes afectados en este gen presentó FM severa (Figura 21, B). Es importante mencionar que los pacientes representados por el complejo *FANCI-FANCD2* y “río abajo” de la vía FA-BRCA (tres en total) sólo se encontraron en los grupos sin FM y con FM severa, respectivamente. Este patrón también se observó para los pacientes con genotipo *FANCB*.

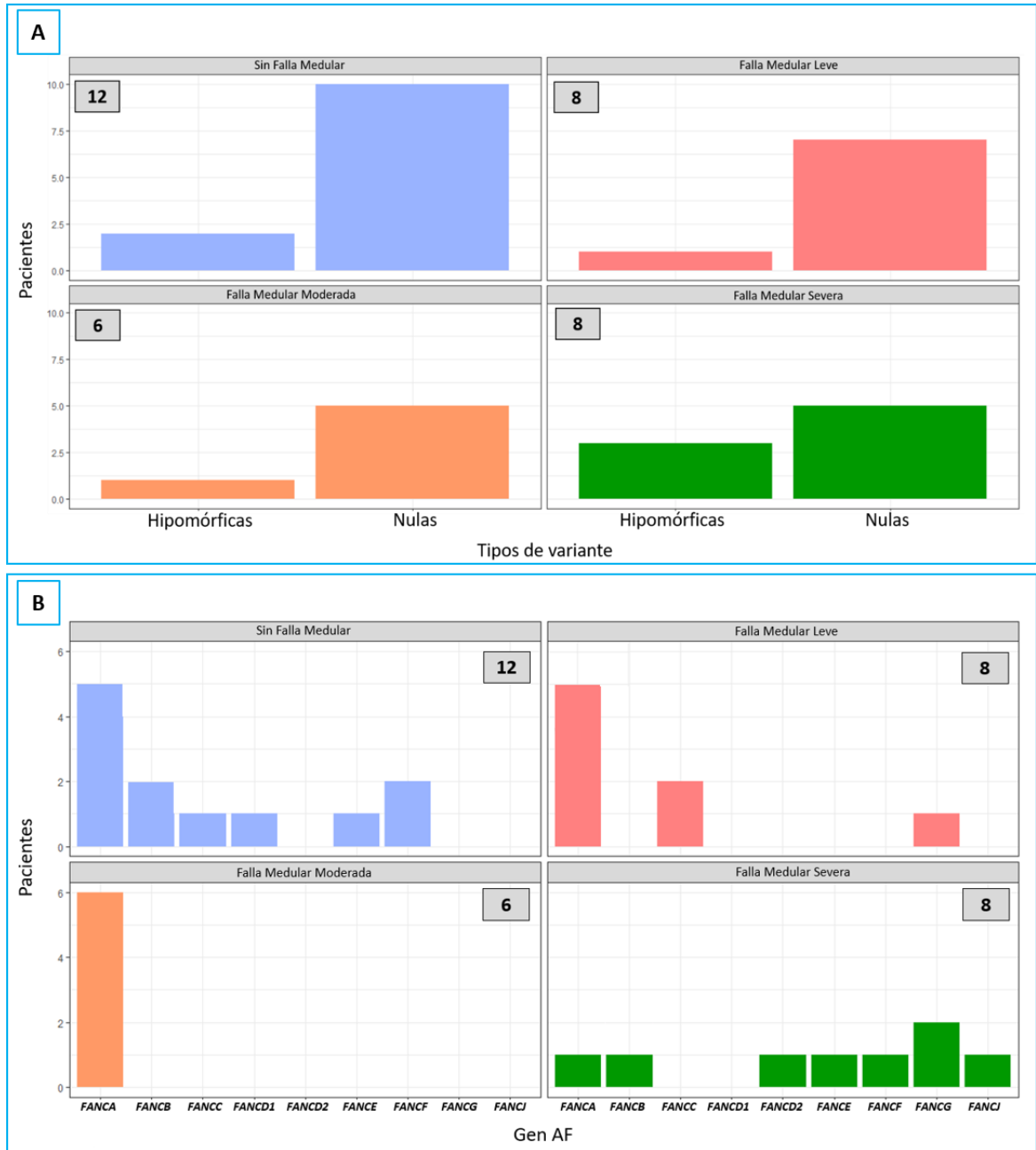


Figura 21. Asociación de la severidad de la falla medular en 34 pacientes con Anemia de Fanconi, con el tipo de variante y genotipo. A. Clasificación de pacientes, de acuerdo al tipo de variante (nulas vs hipomórficas). B. Distribución de pacientes, de acuerdo al genotipo. El número en la parte superior de cada panel corresponde al total de pacientes por grupo.

Debido a esta distribución de frecuencias, decidimos investigar que sucede durante la evolución de la FM desde sus etapas iniciales (leve y moderada) hasta la forma severa, al contrastar los pacientes con genotipo *FANCA* con aquellos que presentaron un genotipo distinto (etiquetado como grupo No-*FANCA*).

En la Figura 22 se observa que los pacientes sin FM presentaron una proporción similar entre ambos grupos, que posteriormente muestra una fuerte tendencia a cambiar hacia *FANCA* en las fases leve y moderada (Prueba exacta de Fisher con $\alpha=0.05$, $p=0.1054$). Sin embargo, y de manera notoria, esta predominancia de *FANCA* disminuye significativamente en la forma severa (Prueba exacta de Fisher con $\alpha=0.05$, $p=0.0062$).

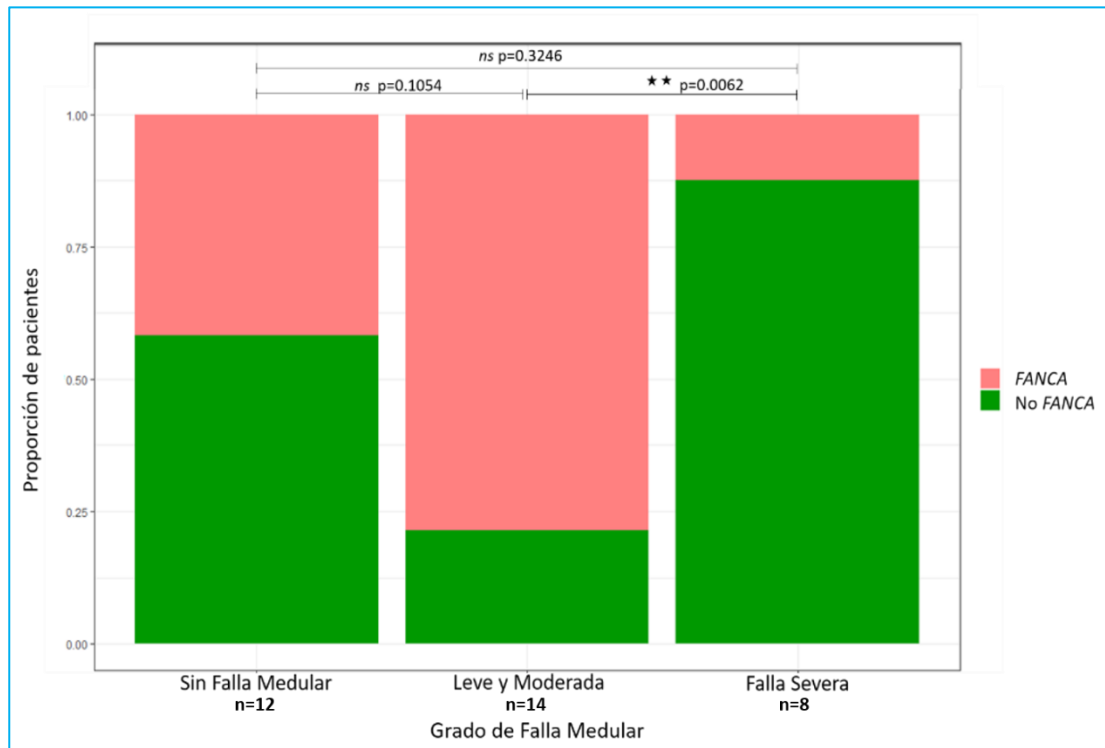


Figura 22: Comparación entre el grado de severidad de la falla medular, luego del agrupamiento de genotipos (*FANCA* vs No *FANCA*). En la Figura: ns=no significativo, de acuerdo a la prueba exacta de Fisher; n= número de pacientes por grupo.

9.5 Resultados que cubren el objetivo 4: Correlacionar el tipo y localización de las variantes encontradas con la proporción y el daño al DNA en las células CD34+.

Para poner en contexto las deficiencias descritas en las CTPH de los pacientes con AF, se obtuvieron simultáneamente muestras de SP y MO en un grupo de 12 pacientes (seis estudiados en el presente trabajo y seis más en un trabajo previo) (Garza-Mayén, 2018). Estas muestras fueron inicialmente contrastadas con individuos sanos (etiquetado como grupo No-AF). El análisis de estas muestras nos permitió evaluar determinadas características de la hematopoyesis en ellos, mediante ensayos de citometría de flujo: la capacidad proliferativa de las CTPH (a través del marcaje con anticuerpo anti-CD34) y respuesta al daño genómico (a través del marcaje con anticuerpo anti- γ H2AX).

De acuerdo a su nivel de FM, los 12 pacientes se distribuyeron en tres categorías: sin FM (dos), FM leve (tres) y FM severa (siete). A partir de este grupo de estudio, se pudo obtener el genotipo de siete pacientes (ver Tabla 12). A causa del tamaño limitado de las muestras con genotipo, decidimos analizar la correlación con el fenotipo celular considerando a los pacientes sin FM y con FM leve como un sólo grupo (etiquetado como “SFM/L”), el cual fue contrastado con los pacientes con FM severa. Por otra parte, también optamos por contrastar los casos con alteraciones en *FANCA* contra aquellos afectados en otros genes de la vía FA-BRCA (etiquetado como *No FANCA*), de manera similar a como se describió en las secciones previas.

Tabla 12. Grupo de pacientes con Anemia de Fanconi, con determinación de genotipo y análisis de CTPH a través de citometría de flujo.

Paciente con AF	Gen alterado	Nivel de Falla Medular
FANC3	<i>FANCA</i>	Leve
FANC24	<i>FANCA</i>	Sin Falla Medular
FANC150	<i>FANCA</i>	Leve
FANC151	<i>FANCA</i>	Sin Falla Medular
FANC32	<i>FANCG</i>	Leve
FANC143	<i>FANCG</i>	Severa
FANC31	<i>FANCD2</i>	Severa

Inicialmente, en este estudio se compararon los tejidos evaluados entre los pacientes con AF y los individuos del grupo No-AF. Se observa que el porcentaje de células CD34+ en SP tiende a ser más abundante en los pacientes con AF, aunque estadísticamente no hubo diferencias (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.121$) (Figura 23, A). Esta situación es opuesta en la MO, ya que la cantidad de células CD34+ fue significativamente inferior en los pacientes (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.025$). Si se considera sólo a los pacientes con AF, no se aprecian diferencias entre ambos tejidos en cuanto a la cantidad de estas células (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.2735$). La distribución de las mismas en los tejidos no presentó variaciones vinculadas a los niveles de FM.

Con relación al gen afectado, es importante destacar que, aunque el grupo estudiado fue reducido (7 pacientes), los individuos con FM severa sólo se encuentran en el grupo *No-FANCA*. Adicionalmente, se observa una mayor cantidad de células CD34+ en SP en los casos en donde el genotipo difería del grupo con *FANCA* alterado (Figura 23, B).

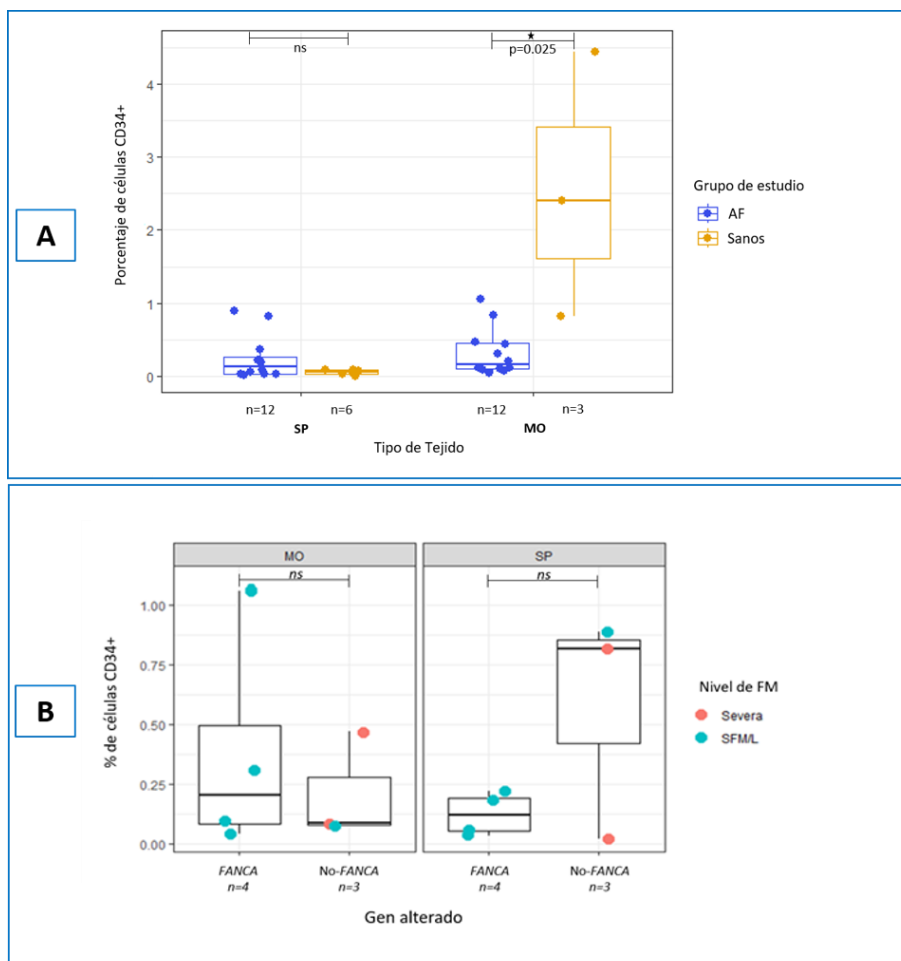


Figura 23. Análisis comparativo de la proporción de células CD34+, de acuerdo al tejido evaluado. A. Proporción de células CD34+, entre pacientes con Anemia de Fanconi (AF) y un grupo de individuos sanos. B. Proporción de células CD34+ en pacientes con AF, de acuerdo al genotipo y el nivel de falla medular. En la Figura: FM: falla medular, MO: médula ósea, n: número de pacientes por grupo contrastado, ns: diferencia no significativa entre grupos, SFM/L: pacientes sin falla medular o con falla medular leve. SP: sangre periférica.

Posteriormente, se evaluó la cantidad de daño en las células CD34+, empleando un anticuerpo anti- γ H2AX, que señala rupturas de doble hebra. Como parámetro de medición se utilizó la intensidad media de fluorescencia (abreviada en inglés como MFI). En la Figura 24, A; se observa que las células CD34+ de los pacientes con AF presentan una cantidad sustancialmente mayor de daño no reparado, en comparación con los individuos sanos (No-AF). Esta diferencia fue significativa en SP (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.005$).

Asimismo, en los pacientes con AF no hubo diferencias entre ambos tejidos en cuanto a la cantidad de daño (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.374$). Sin embargo, en SP se presenta una tendencia a tener mayor daño genómico cuando la FM es severa (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.14$). (Figura 24, B). Respecto al grupo de pacientes con genotipo, se observa que la media de daño es mayor en los

pacientes con genotipo *FANCA* en ambos tejidos, aunque no existe una asociación evidente con el nivel de FM (Figura 24, C).

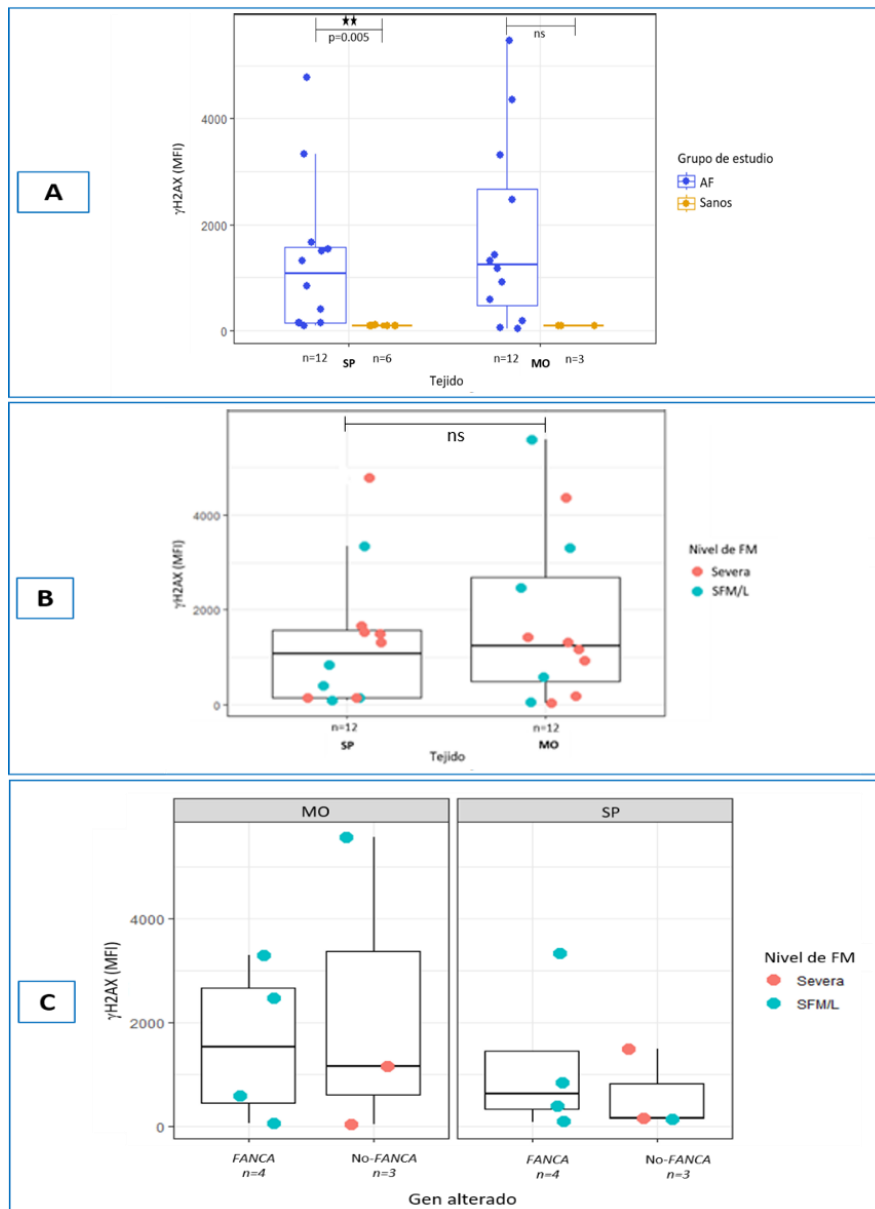


Figura 24. Evaluación del daño genómico en células CD34+, de acuerdo al tejido evaluado. La cantidad de daño se expresa como intensidad media de fluorescencia (MFI) del marcador γ H2AX. A. Daño genómico, entre pacientes con Anemia de Fanconi (AF) e individuos sanos. B. Daño genómico en pacientes con AF, de acuerdo al nivel de falla medular. C. Daño genómico, de acuerdo al genotipo y el nivel de la falla medular. En la Figura: FM: falla medular, MO: médula ósea, n: número de pacientes por grupo contrastado, ns: diferencia no significativa entre grupos, SFM/L: pacientes sin falla medular o con falla medular leve. SP: sangre periférica.

Por otra parte, las células que se encuentran en SP han transitado por un proceso de maduración, que se inicia a partir de la diferenciación de las CTPH, por lo que eventualmente pierden el marcador de superficie CD34. Por ello, las células diferenciadas son consideradas posteriormente como CD34 negativas (CD34-). Decidimos analizar si existían diferencias entre el daño genómico que se encontraba en las células CD34+ y las vs CD34- en los pacientes con AF. En la Figura 25, A; podemos apreciar que tanto en SP como en MO el daño en las células CD34+ es mayor respecto a las células que iniciaron su diferenciación, aunque este resultado no fue estadísticamente significativo (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.23$ para SP y $p=0.11$ para MO). No se observaron diferencias entre ambos tejidos en cuanto a la cantidad de daño (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.25$). Sin embargo, en ambos tejidos los valores más elevados de daño se presentan en los pacientes con FM severa, con respecto al grupo SFM/L (Figura 25, B), siendo esta tendencia muy cercana a la significancia estadística en SP (Prueba de U de Mann Whitney con $\alpha=0.05$, $p=0.052$). Por otra parte, en el grupo de pacientes con genotipo, se observa que la media de daño en sus células diferenciadas es mayor en los pacientes con genotipo *FANCA* en ambos tejidos, aunque no se presentó una asociación evidente con el nivel de FM presentado (Figura 25, C).

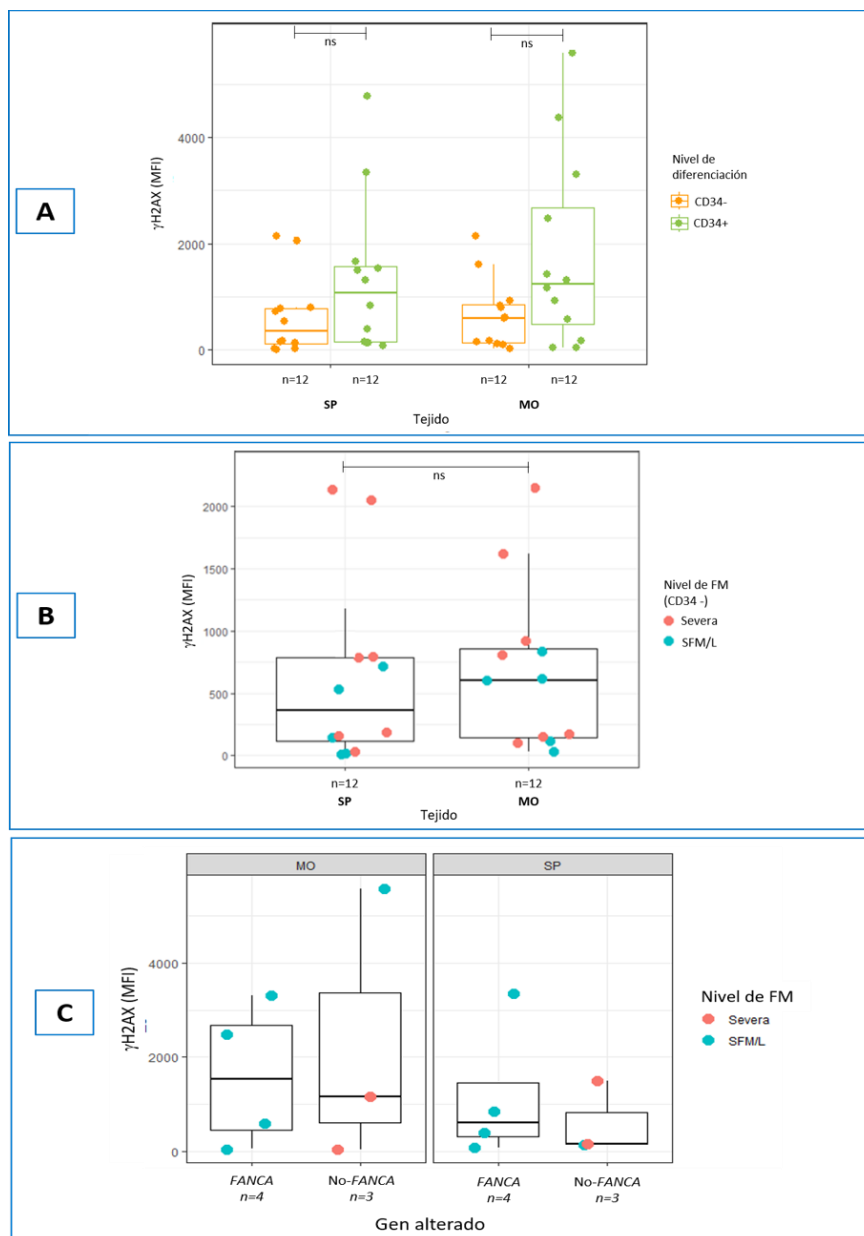


Figura 25. Evaluación del daño genómico, entre células no diferenciadas (CD34+) y células diferenciadas (CD34-), en pacientes con Anemia de Fanconi (AF), de acuerdo al tejido evaluado. A. Daño genómico, de acuerdo a la presencia o ausencia del marcador CD34. B. Daño genómico en células CD34-, de acuerdo al nivel de falla medular. C. Daño genómico en células CD34-, de acuerdo al genotipo y el nivel de falla medular. En la Figura: FM: falla medular, MO: médula ósea, n: número de pacientes por grupo contrastado, ns: diferencia no significativa entre grupos, SFM/L: pacientes sin falla medular o con falla medular leve. SP: sangre periférica.

10. DISCUSIÓN.

Determinación del tipo y localización de las variantes patogénicas en los genes FANC.

Este trabajo es el primero en llevar a cabo una estrategia de genotipificación en un grupo extenso de pacientes mexicanos con AF (58 pacientes). Debido a que la determinación del genotipo en estos pacientes es compleja e implica el uso de diversas metodologías; y que la cantidad de DNA que se puede obtener de las muestras sanguíneas de los pacientes con AF está relacionada al nivel de agravamiento de la falla medular, hubo casos en nuestro estudio en los que el proceso de genotipificación no pudo proseguir hasta la obtención del gen *FANC* afectado, limitando la eficiencia diagnóstica de la estrategia experimental propuesta. Derivado de este proceso, se obtuvo la(s) variante(s) asociada(s) a la enfermedad en 41 de ellos (70.68% del grupo original), provenientes de 33 familias. Un aspecto relevante de este rendimiento es que superó la proyección planteada para la duración del proyecto (\approx 30 pacientes en cuatro años), lo que también refleja la optimización en la captación de pacientes con AF para la búsqueda de genotipo en el Laboratorio de Citogenética del Instituto Nacional de Pediatría.

A pesar de que existen algunas diferencias en cuanto a la proporción de casos en ciertos grupos étnicos, las variantes patogénicas en *FANCA* usualmente son responsables de la enfermedad en cerca de 60% de los individuos afectados (Kimble et al, 2018; Kennedy y D'Andrea, 2005). En este estudio, 56% de los casos presentaron variantes en este gen, frecuencia muy cercana a lo descrito a nivel global. Las deleciones grandes, que se detectaron en todos los casos mediante MLPA; constituyeron el 20.93% de los alelos afectados en *FANCA*, aproximándose también a lo reportado. Este resultado apoya el uso del MLPA como paso inicial en el proceso de genotipificación, considerando representan aproximadamente un cuarto de los pacientes con este genotipo (26.08%).

La deleción más frecuente en los pacientes (*FANCA*: c.(?_-42)_(283+1_284-1)del), afecta los primeros tres exones. Se ha reportado que cerca de la mitad de las deleciones en el gen se prolongan más allá de su extensión, ocasionando la eliminación de genes aledaños y posiblemente afectando la heterogeneidad fenotípica de los pacientes. Asimismo, una elevada proporción de estos eventos (75% aproximadamente) son mediados por recombinación homóloga no alélica entre elementos *Alu-Alu*. Este mecanismo es uno de los más recurrentes en eventos de recombinación intracromosómica en enfermedades causadas por deleciones (Flynn et al, 2014).

La presencia de pequeñas deleciones y cambios puntuales de diverso impacto también son comunes en *FANCA*. Respecto a estos cambios, la variante más frecuente fue *FANCA*: c.2535_2536delCT (identificada en siete de 46 alelos afectados, 15.21%), la cual estuvo presente en cuatro familias, procedentes de tres estados de México (Estado de México, Morelos y Veracruz). De manera interesante, esta variante, ubicada en el exón 27, y que conduce a un corrimiento en el marco de lectura, forma parte del panel de tamizaje de VP en población brasileña con AF, en la que se busca a través de PCR-ARMS (\approx 3% de los casos) (Pilonetto et al, 2017). Debido a estos hallazgos, la búsqueda intencionada de esta variante podría optimizar las primeras etapas de la estrategia de

tamizaje futura de la AF en población mexicana. En contraste, la variante *FANCA*: c.3788_3790del, ubicada en el exón 38 y reportada cerca de 300 veces en la base de datos de AF, se presentó en un sólo caso en nuestro grupo, en estado homocigoto.

El segundo gen más afectado en nuestra cohorte de pacientes con AF fue *FANCB* (cuatro casos, 9.75%). Esta proporción resulta llamativa, ya que su frecuencia global es cercana al 2%. Las VP en este gen están relacionadas a un fenotipo severo, además de presentar una fuerte correlación con la asociación VACTERL-H (Jung et al, 2020, Fiesco-Roa et al, 2019). Existe poca información sobre el rol de la proteína *FANCB* durante la conformación del complejo central. Sin embargo, muchas de las variantes descritas se han vinculado a la disminución de la monoubiquitinación de *FANCI* en modelos *in vitro*, y se ha sugerido que su proximidad con *FANCL* también refleja su función clave para la activación del complejo *FANCD2-FANCI* (Jung et al, 2020). Dos de las tres variantes encontradas en este estudio afectan el extremo N- terminal, que contiene un motivo “ β -hélice” que le permite interactuar con los componentes del complejo CEF (Jung et al, 2020).

Los pacientes con VP en *FANCC* y *FANCG* en nuestra cohorte tuvieron una distribución equitativa (3 pacientes por gen, 7.31%). A nivel mundial, *FANCC* y *FANCG* contabilizan de manera combinada cerca del 20% de los casos con AF, siendo los que presentan mayor número de variantes únicas reportadas después de *FANCA* (Alter et al, 2018). Sin embargo, en nuestro análisis, una sólo variante por cada uno de estos genes fue la responsable del fenotipo de los individuos afectados (*FANCC*: c.67delG, caso familiar) y *FANCG*: c.511-3_511-2delCA (3 pacientes), respectivamente). Esta escasa heterogeneidad de VP y el posterior análisis de las familias afectadas sugieren que la prevalencia de estos genotipos en México está influenciada por variantes con efecto fundador (ver discusión detallada más adelante). Existen estudios poblacionales en los que estos genotipos se encuentran poco representados, entre los que podemos mencionar *FANCC* en India (1.5%), China (5.6%) y Brasil (4.31); así como *FANCG* en Estados Unidos (<5%) y España (6.4%) (Joshi et al, 2023, Nie et al, 2020, Pilonetto et al, 2017, Altintas et al, 2022, Bogliolo et al, 2019).

Otro gen en el que la proporción de pacientes en este estudio difiere con lo reportado es *FANCF* (2%-3% del total de pacientes con AF). En *FANCF* se han reportado poco más de 30 variantes hasta la fecha; y está formado por un sólo exón que codifica para una proteína con un dominio funcional que prácticamente ocupa su totalidad (Zareifar et al, 2019). Las dos variantes encontradas (una que afecta el codón de inicio y la otra de corrimiento de marco de lectura), no han sido reportadas e indican la pérdida de función de la proteína. Las deleciones son el tipo de variante más frecuente en este gen, siendo la variante c.484_485delCT la mejor descrita, asociada a un fenotipo severo. Curiosamente, uno de los pacientes reportados con la variante mencionada también presentó otra alteración en el codón de inicio (*FANCF*:c.2T>C), acompañada de una duplicación 1q (Tryon et al, 2016).

El resto de los genotipos encontrados en este estudio correspondieron a casos únicos (i.e un individuo afectado por cada gen; para *FANCD1*, *FANCD2*, *FANCI*) o a sólo dos casos (*FANCF*). La variante encontrada en *FANCI*, así como una de las presentes en *FANCF*, coinciden con las más reportadas para estos genes a nivel global. Considerados de manera individual, estos genes también presentan frecuencias bajas a nivel global (2%-4%) (Niraj et al, 2018). Sin embargo, es importante mencionar que todos los pacientes del grupo referido, salvo *FANCF*, se consideran como genes clásicos de predisposición a cáncer hereditario de mama y ovario, además de estar frecuentemente asociados a anomalías físicas (Fiesco-Roa et al, 2019). Es por ello que, en estos casos, se recomienda extender el análisis molecular a los familiares, a fin de encontrar individuos portadores en riesgo. Por otra parte, se requiere que estos genotipos estén representados por un número mayor de individuos para poder establecer correlaciones genotipo-fenotipo más robustas.

Análisis de variantes con efecto fundador.

Durante este proceso de genotipificación se encontraron dos variantes con efecto fundador. Los tres pacientes con una VP en *FANCC* representan un caso familiar de AF, de ascendencia Menonita, procedentes del Estado de Tamaulipas. En ellos se anticipó el genotipo a partir de la revisión de la historia migratoria de los antepasados de su comunidad, lo cual resalta la gran utilidad de esta información en estudios dirigidos.

Los asentamientos Menonitas en América del Norte señalan su arribo a Canadá previo a su llegada a México, a principios del siglo XX. La variante encontrada (*FANCC*: c.67delG) había sido reportada con anterioridad en individuos holandeses con AF del mismo grupo étnico. La confirmación citogenética de la enfermedad en el caso índice permitió extender el ensayo molecular por secuenciación de Sanger en sus hermanos, encontrándose 2 afectados y 6 portadores. La corroboración del efecto fundador, a través de la búsqueda de haplotipos, utilizando SNVs previamente reportados y aledaños a la variante, constituye un enfoque metodológico que ha sido ampliamente replicado para validar otras VP similares en diversas poblaciones con AF (Tipping et al, 2001; Park et al, 2015, entre otros).

Este tipo de diseño contrasta con la validación del efecto fundador en los tres pacientes del grupo étnico *Mixe*, debido a que la variante descrita en *FANCG* era desconocida. Sin embargo, ejemplifica muy bien la combinación de metodologías moleculares para la caracterización de una variante, ya que implicó el hallazgo de una delección pequeña en *FANCG* por SNG, corroborada por secuenciación de Sanger. Luego de evaluar su efecto patogénico sobre el mecanismo de corte y empalme a través de herramientas *in silico*, surgió la presunción de una posible relación entre los pacientes, a partir de la revisión de los expedientes (procedencia, apellidos, mismo grupo étnico), razón por la cual se realizaron microarreglos de alta densidad. La información obtenida desde esta plataforma se utilizó para construir una inferencia que posteriormente fue validada: los pacientes compartían haplotipos en una zona del cromosoma 9 donde se localiza *FANCG*, evidenciando su efecto fundador, mientras que el análisis de ancestría

puso en contexto la reciprocidad entre los pacientes *Mixes* y los marcadores genéticos que distinguen a la población mexicana.

La búsqueda y corroboración de las variantes de efecto fundador tiene un gran impacto social, ya que establece precedentes para la planificación de programas de tamizaje diseñados para la búsqueda de portadores y afectados por la AF en regiones específicas, a fin de proporcionar a las familias afectadas un consejo genético preciso y establecer la verdadera frecuencia de estos alelos en poblaciones en riesgo.

Variantes encontradas en pacientes con Anemia de Fanconi que desarrollaron cáncer.

La principal causa de mortalidad en la AF, en adición a la progresión de la FM, es el cáncer. En nuestro grupo, tres pacientes (7.31%) presentaron manifestaciones asociadas a neoplasias. Esta proporción es inferior a lo reportado en la enfermedad, donde se ha estimado una prevalencia cercana al 13%-15% en aquellos casos que no han recibido trasplante de MO, incrementándose hasta 24%-30% en edad adulta (Alter et al, 2018; Altintas et al, 2022). Dos pacientes presentaron VP en *FANCA*, mientras que el caso restante corresponde a un individuo con VP en *FANCD2*.

Los dos pacientes con VP en *FANCA* desarrollaron manifestaciones típicamente asociadas a malignización hematológica en la AF: uno con LMA y otro con SMD, relacionadas con alteraciones citogenéticas clonales en el cromosoma 7. En ambos casos se hallaron VP no reportadas, en estado homocigoto. La variante observada en el paciente con LMA (*FANCA*: c.1318G>T, p.Glu440Ter), también se encontró en otro paciente de esta misma cohorte, en estado homocigoto, aunque este último hasta la fecha no ha presentado malignización. Dicho resultado discrepa con la propuesta que sugiere una alta incidencia de cáncer en pacientes con VP nulas en *FANCA* (Castella et al, 2011). Algunos autores señalan que la expresión disminuida y la deficiencia de la proteína en casos de LMA esporádico reflejan su impacto negativo en el desarrollo de esta neoplasia (Tischkowitz et al, 2004).

El análisis *in silico* de la variante reportada en el paciente con SMD (*FANCA*: c.1471-7T>G) indicó la afectación de un sitio regulador de corte y empalme en el intrón 15 (*branch point*). Aunque las alteraciones en este proceso se han asociado a retraso en el desarrollo y anomalías del sistema nervioso en la AF, un estudio en Israel señala que los pacientes con este tipo de variantes desarrollan la primera neoplasia a edades más tempranas que aquellos que presentan deleciones, independientemente del gen afectado (Steinberg-Schemer et al, 2020). En este caso en particular, la evolución clonal fue rápida, y el paciente falleció al poco tiempo de la determinación del SMD.

El paciente afectado en *FANCD1* (*BRCA2*) desarrolló un meduloblastoma a los cuatro años de edad. Esta aparición temprana de neoplasias se asemeja a lo descrito para este genotipo, que presenta una probabilidad de 97% de desarrollar algún cáncer para los 5 años (Alter et al, 2007). Aunque el tumor encontrado es frecuente en niños, presenta una correlación genotipo-fenotipo sólida en la AF, encontrándose casi exclusivamente en pacientes con VP en este gen, o en su gen asociado (*FANCN/PALB2*) (McReynolds et al, 2021). En el paciente, cuyos antecedentes familiares estaban asociados a cáncer (una de

sus hermanas falleció de leucemia linfoblástica aguda), se encontraron dos VP distintas (ambas de corrimiento de marco de lectura). Una de ellas (*FANCD2*: c.658_659del) se ha reportado frecuentemente en este tumor embrionario, especialmente en casos familiares (Miele et al, 2015). Por otra parte, se ha descrito que dos variantes de efecto nulo (como en el paciente mencionado) no manifiestan formas más tempranas o agresivas de cáncer, comparadas con aquellos casos en los que se presentan variantes hipomórficas (McReynolds et al, 2021). Estos resultados respaldan la necesidad de comenzar tempranamente la vigilancia oncológica en pacientes con AF.

Caracterización de la patogenicidad de las variantes mediante análisis bioinformático.

Durante el análisis de variantes, es necesario considerar la posición en la que ocurren (región codificante o no codificante) y el número de nucleótidos afectados. En base a esta clasificación, y debido a que la caracterización clínica y experimental de su efecto sólo es posible en un número limitado de casos, se han desarrollado muchos programas computacionales para estimar sus implicaciones. Actualmente, estas aplicaciones son entrenadas y validadas de acuerdo a datos experimentales, respaldando su uso para la interpretación acertada de variantes cuando son empleadas en el contexto adecuado (Katsonis et al, 2021).

La valoración del impacto de las variantes no sinónimas sobre una proteína debe considerar múltiples características, como capacidad de plegamiento, interacciones, modificaciones post-traduccionales, solubilidad, entre otras. Por esta razón, existen múltiples métodos *in silico* para estudiarlas. Sin embargo, estos métodos están limitados a la disponibilidad de información estructural de las proteínas de interés. En tal sentido, el aislamiento experimental de las proteínas de la vía FA/BRCA de forma individual ha resultado desafiante hasta la fecha, sobre todo de aquellas que conforman el complejo central, debido a la elevada dificultad metodológica para el estudio de sus interacciones. Técnicas como la crio-microscopía electrónica y la espectrometría de masas han permitido el refinamiento tridimensional parcial de algunos dominios a gran resolución (Shakel et al, 2019; Wang et al, 2019). En consecuencia, bases de datos como PDB (*Protein Data Bank*) o Uniprot carecen de la estructura cristalográfica completa de muchas de ellas. Ante este panorama, el uso de predictores de alta calidad, como AlphaFold o RosettaFold (basadas en algoritmos de aprendizaje profundo ("*deep learning*")) resulta muy útil para el modelamiento de estas proteínas (Katsonis et al, 2021).

En este estudio se utilizó AlphaFold, en adición a otros programas de visualización, como Dynamut y Pymol, para obtener una estructura tridimensional más precisa del impacto de las variantes no sinónimas de sentido erróneo sobre las proteínas silvestres. Un ejemplo de este enfoque fue la caracterización de la variante *FANCA* c.1874G>C (p.Cys625Ser), originalmente clasificada como variante de significado incierto; de la cual evaluamos su impacto deletéreo sobre la estabilidad e interacciones moleculares. Se utilizaron de forma complementaria otros programas, basados en patrones evolutivos, funcionales y de homología, para disponer de una visión integrativa del efecto de estos cambios.

De manera similar, se procedió con la evaluación de las variantes sin sentido y las de corrimiento de marco de lectura. Al respecto, es importante mencionar que no todas ocasionan pérdida o ganancia de función, ya que un grupo de estas variantes pueden ser polimórficas o afectar la sección final de las proteínas, evitándose la degradación del mensajero (Hu y Ng, 2012); por lo que su evaluación bioinformática también es necesaria. A partir de la predicción de los programas utilizados en nuestro estudio, se concluyó que las variantes de este tipo analizadas afectan la función de las proteínas.

Para la caracterización de variantes ubicadas en zonas no codificantes, los programas que empleamos permitieron examinar modificaciones en sitios considerados como canónicos del proceso de corte y empalme, ya sea en sitios aceptores o donadores (como en el caso de la variante fundadora *FANCG*:c.511-3_511-2delCA); o en puntos de ramificación (como en *FANCA* (c.1471-7T>G)). El resultado en estas alteraciones puede ocasionar la omisión de un exón o la generación de un nuevo sitio críptico. Las variantes ubicadas en los dos primeros residuos aledaños a los bordes intrón-exón (por ejemplo, la descrita en *FANCG*) tienen una evidencia de patogenicidad muy alta, de acuerdo con las guías de la ACMG (Richards et al, 2015). Por otra parte, los puntos de ramificación son fundamentales para la ubicación del complejo de corte y empalme (*spliceosome*). Pueden ubicarse entre -9 y -400 pares de bases río abajo del sitio aceptor, y están formados por secuencias altamente degeneradas, lo cual dificulta su localización precisa (Riolo et al, 2021). De acuerdo al análisis computacional de la variante que afecta uno de estos sitios, este cambio tiene una alta probabilidad de perturbar su reconocimiento, que podría ocasionar la pérdida de afinidad de las proteínas snRNP.

De las variantes descritas en este trabajo, cerca del 80% fueron clasificadas como patogénicas o probablemente patogénicas, mientras que el resto fueron asignadas como variantes de significado incierto (VUS, por su abreviatura en inglés). De manera peculiar, algunas de estas VUS se encuentran reportadas en la base de datos de pacientes con AF (FADB), o tienen interpretaciones conflictivas de patogenicidad en la base de datos de ClinVar. En virtud de esta situación, la validación experimental de las variantes, en especial de aquellas catalogadas como VUS, sustentadas previamente por estudios *in silico* robustos, es fundamental para la resolución final de la patogenicidad final de las mismas.

Búsqueda de una correlación genotipo-fenotipo.

Existen numerosos estudios sobre la frecuencia de los genotipos *FANC* en diversas poblaciones. Sin embargo, cada vez es más frecuente la búsqueda de correlaciones entre el gen afectado y la presentación fenotípica de los pacientes. Esta tendencia es más evidente en reportes en donde se documentan descripciones físicas y clínicas más detalladas, en un número de casos considerable (Altintas et al, 2022; Alter et al, 2018). En nuestro grupo de pacientes con genotipo, dichas asociaciones se analizaron en dos de los tres principales grupos de manifestaciones de la enfermedad (alteraciones en el desarrollo físico y hematológico), ya que sólo tres de los 41 pacientes evidenciaron procesos neoplásicos.

Debido a la distribución asimétrica de los genotipos en los pacientes (56% *FANCA* vs 44% en los otros 8 genes *FANC*), la búsqueda de las correlaciones genotipo-fenotipo implicó el agrupamiento de los pacientes en base a parámetros distintos a la comparación entre genes, como el tipo de variante, la participación en complejos que forman parte de la vía FA-BRCA, y la diferencia entre individuos afectados en *FANCA* respecto al resto de los genotipos. Este tipo de evaluaciones se justifica, ya que el establecimiento de correlaciones con el fenotipo facilita la estratificación de grupos de riesgo (Altintas et al, 2022).

La descripción del fenotipo físico en pacientes con AF usualmente integra las características incluidas en la asociación VACTERL-H y en el acrónimo PHENOS. En este proyecto se observó una proporción elevada de pacientes con anomalías en la pigmentación de la piel, siendo el rasgo encontrado con mayor frecuencia (65.85%). Estas alteraciones se asociaron significativamente con la presencia de variantes nulas. En un estudio realizado en 279 pacientes con AF procedentes de la India, se encontró una frecuencia similar (63.44%), con manifestaciones dermatológicas clásicamente descritas en la AF como manchas *café au lait* e hiperpigmentación generalizada (Shukla et al, 2018). La correlación de esta característica con variantes nulas en nuestro grupo es llamativa, ya que no coincide con lo descrito en otros estudios más amplios (Fiesco-Roa et al, 2019), por lo que resulta necesario precisar si esta asociación se mantendrá cuando el grupo de pacientes mexicanos con AF se incremente.

Por otra parte, las alteraciones en las extremidades superiores también se presentaron en una proporción importante (56%), seguidas de las asociadas a talla baja (44%), aunque no se relacionaron con ninguna asociación en particular. En conjunto, estas tres características forman parte de las más reportadas en los pacientes, y son consideradas como distintivas en la AF (27%-43%) (Fiesco-Roa et al, 2019). Aun cuando se ha estimado que la frecuencia de la asociación VACTERL-H en pacientes con AF oscila entre 5%-33% (Fiesco-Roa et al, 2019); en nuestro grupo fue ligeramente menor al límite inferior mencionado (4.87%). Sin embargo, esta situación puede deberse a que la búsqueda de anomalías vertebrales y del eje radial, así como las renales y cardíacas, son dependientes de procedimientos de imagen, que no se realizan de manera rutinaria en los pacientes. Por ello, es importante que estas evaluaciones sean incorporadas de forma sistemática como parte de su valoración dismorfológica.

A pesar de que no se contrastaron directamente los fenotipos de pacientes considerando sólo el gen *FANC* (por ejemplo, *FANCA* vs *FANCB*), algunas de las características descritas se vinculan a lo reportado para estos genes en la literatura. En particular, las VP en *FANCB* están relacionadas a un fenotipo severo, además de presentar una fuerte correlación con la asociación VACTERL-H (Jung et al, 2020, Fiesco-Roa et al, 2019). En nuestros pacientes con este genotipo, uno de ellos cumplió criterios para esta asociación ($\geq 3/8$), mientras que otro cumplió criterios para PHENOS ($\geq 4/6$). En adición, la mitad de los casos de este grupo fueron diagnosticados con AF antes de los seis meses de edad, muy probablemente debido a sus manifestaciones físicas. Estos datos coinciden con la corta expectativa de vida y la severidad del fenotipo característica de los pacientes afectados en este gen (Fiesco-Roa et al, 2019). Por otra parte, uno de los tres pacientes

afectados en *FANCF*, también presentó la asociación VACTERL-H. A pesar de que los reportes que incluyen la descripción de los pacientes afectados en este gen son muy limitados (reflejo de ello es que no se encuentra representado en el análisis genotipo-fenotipo más grande hecho hasta ahora para la AF (Fiesco-Roa et al, 2019)), un reporte señala que 30% de estos casos desarrollan la asociación VACTERL, aunque sin hidrocefalia (Faivre et al, 2005). Finalmente, aunque algunos pacientes comparten la misma variante y pertenecen a familias no relacionadas (por ejemplo, los pacientes afectados en *FANCG*), el espectro de manifestaciones físicas no fue homogéneo, de manera similar a como se ha reportados en otros casos (Jung et al, 2020).

En lo concerniente al análisis hematológico, 64.70% de los pacientes tuvo datos de falla medular al momento del diagnóstico de la AF. Este porcentaje difiere de lo reportado en la literatura ($\approx 70\%$ -80%) (Dufour, 2017). No obstante, cabe considerar que el promedio de edad del grupo estudiado fue inferior al promedio de diagnóstico de la FM en la enfermedad (6.09 vs 7). Por otra parte, respecto a la clasificación de la severidad de la FM, sobresale la predominancia del genotipo *FANCA* en casos de FM leve o moderada, lo cual contrasta significativamente con la diversidad de genotipos encontrada en la forma severa. Este resultado es semejante a lo reportado en un grupo de 116 pacientes en Brasil, en el que se estudió el grado de FM en pacientes con VP en *FANCA* con *FANCC*, *FANCE* y *FANCG* (Pilonetto et al, 2022). Asimismo, este resultado también nos permite inferir que la evolución de la FM en los otros genotipos puede agravarse con mayor rapidez que en *FANCA*, aunque es necesario obtener más datos para validar esta aseveración.

Correlación del tipo y localización de las variantes con la cantidad y nivel de daño en el DNA en células CD34+.

En pacientes con AF, la determinación del número y cantidad de daño en las CTPH (identificadas a través del marcador CD34+) representa una oportunidad valiosa para estudiar su localización y funcionalidad, en un escenario apegado a lo que ocurre *in vivo*. En tal sentido, la cantidad de CTPH en la MO de los pacientes analizados fue significativamente menor que en individuos sanos. Este resultado es reflejo de su agotamiento para producir células sanguíneas primitivas en el nicho hematopoyético, en el que se ha reportado la sobreexpresión de p53 y de genes vinculados a la vía de TGF- β , que actúan como señales supresoras de crecimiento (Kelly et al, 2007; Garaycochea y Patel, 2014). En contraste, en sangre periférica se detectó una mayor proporción de células CD34+ en los pacientes con AF, respecto a los individuos sanos (aunque no significativa estadísticamente). Este escenario se puede atribuir al desprendimiento de las CTPH desde la MO, ante el déficit de factores de crecimiento y diferenciación propios del microambiente medular pro inflamatorio descrito en la AF (Ceccaldi et al, 2012; Zhang et al, 2015).

La elevada cantidad de células primitivas en circulación en la AF también podría relacionarse con una mayor cantidad de daño genómico (rupturas de doble hebra). Al evaluar este parámetro, aunque se presentó un déficit en la reparación en los dos tejidos evaluados, la diferencia resultó significativa en sangre periférica, con respecto al grupo

control (Figura 24, A). Esto sugiere que la reparación deficiente de estas lesiones, ante la inactivación de la vía FA/BRCA, a través de mecanismos propensos a error (como la unión de extremos no homólogos), puede contribuir en el deterioro de la diferenciación de las CTPH, comprometiendo su funcionalidad y favoreciendo su escape hacia la circulación periférica. Otro factor que puede estar influenciando esta dinámica en los pacientes con AF es la inducción anormal de la expresión de *MYC*, ya que promueve la proliferación de una subpoblación de CTPH ante el daño creciente en el DNA, contrarrestando la acción de p53 (Rodríguez et al, 2022).

Durante el análisis de las células que iniciaron su diferenciación (CD34-), se observó una tendencia del daño genómico a incrementarse en las etapas más avanzadas de la falla medular en sangre periférica. Esta distinción es relevante, ya que nos puede indicar que el nivel de daño no reparado en las células que se mantienen en circulación es paralela a la evolución de la FM en SP. Se ha descrito que una proporción significativa de las células inmaduras son seleccionadas y eliminadas ante la acumulación de daño genómico mediante apoptosis, que produce una depleción de los linajes sanguíneos y modula la progresión de la falla medular (Ceccaldi et al, 2012). Es posible que en los pacientes con falla severa, sus células inmaduras circulantes, que representan aquellas capaces de sostener una hematopoyesis tan mermada, presenten niveles más bajos de la proteína *MYC*, ya que se ha reportado que los niveles de expresión de este gen disminuyen de acuerdo al agravamiento de la anemia aplásica (Rodríguez et al, 2022).

Durante el estudio de las CTPH en los pacientes con AF, no se pudo analizar la relevancia de las VP en los genes *FANCA* sobre los procesos evaluados, debido al reducido número de pacientes con genotipo en el grupo de muestras pareadas (representados en su mayoría por *FANCA*, cuatro de siete). Por lo tanto, los resultados obtenidos fueron de carácter descriptivo y no permitieron establecer correlaciones robustas. Sin embargo, destaca el hecho de que los individuos con VP ubicadas en genes distintos a *FANCA* presentaron en promedio una proporción de células CD34+ en sangre periférica superior a lo observado en el grupo con VP en *FANCA* (Figura 23). Esta situación también podría asociarse con lo observado en el análisis de la severidad medular, ya que el grupo de FM severa está poco representado por individuos del genotipo *FANCA*. No obstante, se requiere de un análisis con mayor número de pacientes, a fin de confirmar esta inferencia.

Comentarios finales.

El desarrollo de este trabajo permitió la identificación de los genotipos más frecuentes (y sus respectivas variantes asociadas) en los pacientes con AF en México. Esta información es muy valiosa para la estructuración de una estrategia de genotipificación adaptada a la casuística del país, con el propósito de aumentar la eficiencia del diagnóstico molecular de la enfermedad, a fin de extenderlo a una mayor cantidad de pacientes y mejorar su factibilidad (tiempo y costo).

Más de la mitad de nuestros pacientes presentó VP en *FANCA*, de manera similar a lo encontrado en otras poblaciones. No obstante, el análisis de la distribución de los individuos afectados en los otros genes *FANCA*, en la que destaca la alta incidencia en

FANCB y el reducido número de casos y variantes con genotipo *FANCC* y *FANCG*, contrasta con lo reportado a nivel mundial y focaliza la importancia del cribado inicial (“*screening*”) en población mexicana, en las que no existen reportes previos de la frecuencia de la AF. Asimismo, parte de este ejercicio implicó el descubrimiento y descripción de variantes de efecto fundador, que merecen nuestra atención debido a su impacto social y porque ejemplifican el uso de estrategias dirigidas que a futuro deben aplicarse en zonas geográficas donde la prevalencia de la enfermedad podría ser muy alta.

A pesar de que no se realizaron ensayos funcionales para validar el efecto de las variantes no descritas en la AF, el uso de herramientas *in silico* y de bases de datos permitió profundizar y ofrecer mayor contexto acerca del potencial efecto patogénico de las mismas, facilitando su clasificación. Este ejercicio facilitó la integración de la información clínica para la búsqueda de correlaciones genotipo-fenotipo. A partir de este análisis se pudo establecer que, para la población estudiada, las alteraciones en la pigmentación de la piel dependen del tipo de variante, y que la evolución de la falla medular correlaciona con la presencia de VP en *FANCA*, al compararlas con los otros genotipos descritos. Finalmente, el análisis de las CTPH de un grupo de pacientes permitió valorar el efecto de una vía FA-BRCA deficiente, que conduce a la acumulación de daño en el DNA, sobre su capacidad de proliferación y diferenciación; observándose una reducción en su producción en la MO, con tendencia a escapar a SP con mayor acumulación de daño no reparado.

En conjunto, estos hallazgos resaltan la importancia de la identificación oportuna de las VP responsables del fenotipo de los pacientes con AF en México para brindar un asesoramiento genético adecuado a las familias, optimizar su manejo clínico (sobre todo para su inclusión en protocolos de trasplante de MO) y revelar ciertas asociaciones con el fenotipo, tanto físico como celular, durante el curso de la enfermedad. De esta forma, el conocimiento de la composición genómica de las familias con AF en el país permitirá una mejor planificación de futuras estrategias de cribado, a fin de optimizar el alcance de la misma al mayor número de pacientes.

11. CONCLUSIONES.

- a. Este es el primer estudio de genotipificación de pacientes con AF en México. La estrategia de genotipificación implementada permitió la detección de las variantes genéticas responsables de la AF en 41 pacientes mexicanos (34 familias). 56% de los casos presentó variantes patogénicas en *FANCA*, seguidos por los afectados en *FANCB* (9.75%) y después las variantes en *FANCC*, *FANCG* y *FANCF*, cuya frecuencia fue equitativa (7.3% cada uno). La distribución de las VP en los genes *FANC* en los pacientes mexicanos, difiere de la reportada a nivel mundial.
- b. Se encontraron 7 tipos de variantes, siendo las más frecuentes las de sentido erróneo y las de cambio en el marco de lectura, que fueron identificadas en la misma proporción (21.95%). La variante con más alelos afectados (7) en este estudio fue *FANCA*: c.2535_2536delCT. Dos variantes en *FANCA*, y una en *FANCD2*, fueron encontradas en pacientes que desarrollaron cáncer.
- c. Se detectaron dos variantes con efecto fundador en nuestra población de pacientes con AF; la encontrada en *FANCC* en comunidad menonita (Tamaulipas) y la localizada en *FANCG*, en comunidad *Mixe* (Oaxaca). Su descripción es fundamental para la búsqueda de individuos portadores (heterocigotos) y afectados en poblaciones aisladas.
- d. En este estudio, las alteraciones físicas más frecuentes correspondieron a anomalías en la pigmentación de la piel (65.85%), anomalías en las extremidades superiores (56%) y talla baja (44%). No se encontraron asociaciones entre genotipo y las alteraciones descritas en VACTERL-H. La frecuencia de los cambios dermatológicos, que forma parte del acrónimo PHENOS, se relaciona significativamente con la presencia de variantes nulas.
- e. Existen diferencias en la severidad de la falla medular en pacientes con genotipo *FANCA*, ya que prevalece en ausencia de FM o en las formas leves y moderadas, cuando es contrastada con pacientes con otros genotipos.
- f. Los pacientes con AF tienen un número reducido de células troncales y progenitoras hematopoyéticas en la médula ósea, comparadas con muestras de individuos sanos. Asimismo, estas células tienen una tendencia a encontrarse incrementadas en sangre periférica, con mayor daño genómico. No se pudo realizar la evaluación del impacto del gen afectado sobre estos hallazgos, razón por la cual es necesario un número mayor de pacientes para poder realizar este tipo de asociaciones.
- g. Finalmente, la identificación oportuna de las variantes responsables del fenotipo en la AF es un prerequisite para el asesoramiento genético de las familias, manejo clínico de los pacientes (en especial para el trasplante de médula ósea), el descubrimiento de correlaciones genotipo-fenotipo y para conocer la composición genómica de esta entidad en la población mexicana.

12. LIMITACIONES Y PERSPECTIVAS.

Limitaciones.

Cuando se usan métodos relacionados con SNG, es posible que la implementación de un panel o de la modalidad de exoma culmine con un resultado negativo, planteando la búsqueda de VP a través de enfoques que impliquen una mayor cobertura génica (como la SNG de genoma completo), que no estuvieron incluidas durante el desarrollo del proyecto.

Debido a la complejidad en la presentación clínica de la enfermedad, algunas características pueden describirse de manera limitada o ignorarse en el expediente si no son buscadas de manera intencionada durante la exploración del paciente. Esto podría dificultar el establecimiento de correlaciones genotipo-fenotipo con mayor precisión.

La evaluación de procesos celulares como la apoptosis, a través de citometría de flujo en pacientes con AF resulta valiosa para complementar el análisis del efecto del daño no reparado sobre las células troncales. Sin embargo, esta estimación no se puede realizar debido al acceso limitado de las muestras y al compromiso hematológico de los pacientes, que condiciona la cantidad de CTPH disponibles. Idealmente, se requiere una mayor cantidad de CTPH de los pacientes, ya que esta evaluación suele implicar el uso de células vivas.

Perspectivas.

Se espera continuar con el proceso de genotipificación, y mejorar la estrategia de obtención de genotipos para alcanzar un mayor porcentaje de pacientes genotipificados, de forma costo-efectiva y rápida, adaptada a la casuística propia de México.

Después de la obtención del genotipo, el proceso de clasificación de variantes puede implicar que algunas de ellas sean catalogadas como “probablemente patogénicas” o como variantes de significado incierto. Las recomendaciones en estos casos abogan por la realización de estudios funcionales. Esta caracterización es muy importante, considerando la gran proporción de variantes que no han sido descritas en otros pacientes o que carecen de análisis moleculares previos. Resultaría importante obtener muestras de RNA de los pacientes con variantes que afectan el proceso de corte y empalme, a fin de validar su impacto sobre la remoción de intrones.

Por otra parte, la obtención de líneas celulares derivadas directamente de los pacientes o que sean obtenidas a partir de mutagénesis dirigida permitiría la realización de ensayos funcionales que evalúen su sensibilidad ante agentes inductores de ICLs, o ensayos de inmunofluorescencia para estudiar la localización nuclear (como en el caso de variantes en *FANCA* que afecten esos dominios). Respecto a las dos poblaciones en las que se detectaron VP fundadoras, es necesario hacer un estudio de genotipificación dirigido en ambas poblaciones, para detectar individuos afectados y portadores y poder dar un consejo genético apropiado.

13. REFERENCIAS.

1. Adamo, A., Collis, S. J., Adelman, C. A., Silva, N., Horejsi, Z., Ward, J. D., Martinez-Perez, E., Boulton, S. J., & La Volpe, A. (2010). Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. *Molecular Cell*, 39(1), 25–35. doi: [10.1016/j.molcel.2010.06.026](https://doi.org/10.1016/j.molcel.2010.06.026)
2. Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. doi: [10.1101/gr.094052.109](https://doi.org/10.1101/gr.094052.109)
3. Alter, B. P. (2014). Fanconi anemia and the development of leukemia. Best Practice & Research. *Clinical Haematology*, 27(3–4), 214–221. doi: [10.1016/j.beha.2014.10.002](https://doi.org/10.1016/j.beha.2014.10.002)
4. Alter, B. P., & Giri, N. (2016). Thinking of VACTERL-H? Rule out Fanconi Anemia according to PHENOS: VACTERL-H, PHENOS, and Fanconi Anemia. *American Journal of Medical Genetics. Part A*, 170(6), 1520–1524. doi: [10.1002/ajmg.a.37637](https://doi.org/10.1002/ajmg.a.37637)
5. Alter, B. P., Giri, N., Savage, S. A., & Rosenberg, P. S. (2018). Cancer in the National Cancer Institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. *Haematologica*, 103(1), 30–39. doi: [10.3324/haematol.2017.178111](https://doi.org/10.3324/haematol.2017.178111)
6. Altintas, B., Giri, N., McReynolds, L. J., Best, A., & Alter, B. P. (2023). Genotype-phenotype and outcome associations in patients with Fanconi anemia: the National Cancer Institute cohort. *Haematologica*, 108(1), 69–82. doi: [10.3324/haematol.2021.279981](https://doi.org/10.3324/haematol.2021.279981)
7. Ameziane, N., Errami, A., Léveillé, F., Fontaine, C., de Vries, Y., van Spaendonk, R. M. L., de Winter, J. P., Pals, G., & Joenje, H. (2008). Genetic subtyping of Fanconi anemia by comprehensive mutation screening. *Human Mutation*, 29(1), 159–166. doi: [10.1002/humu.20625](https://doi.org/10.1002/humu.20625)
8. Auerbach, A. D. (1993). Fanconi anemia diagnosis and the diepoxybutane (DEB) test. *Experimental Hematology*, 21(6), 731–733. <https://pubmed.ncbi.nlm.nih.gov/8500573/>
9. Auerbach, A. D. (2003). Diagnosis of Fanconi anemia by diepoxybutane analysis. *Current Protocols in Human Genetics*, 85(1), 8.7.1-8.7.17. doi: [10.1002/0471142905.hg0807s85](https://doi.org/10.1002/0471142905.hg0807s85)
10. Auerbach, Arleen D. (2009). Fanconi anemia and its diagnosis. *Mutation Research*, 668(1–2), 4–10. doi: [10.1016/j.mrfmmm.2009.01.013](https://doi.org/10.1016/j.mrfmmm.2009.01.013)
11. Barbosa, F. B., Cagnin, N. F., Simioni, M., Farias, A. A., Torres, F. R., Molck, M. C., Araujo, T. K., Gil-Da-Silva-Lopes, V. L., Donadi, E. A., & Simões, A. L. (2017). Ancestry informative marker panel to estimate population stratification using genome-wide human array. *Annals of Human Genetics*, 81(6), 225–233. <https://doi.org/10.1111/ahg.12208>

12. Bogliolo, M., Pujol, R., Aza-Carmona, M., Muñoz-Subirana, N., Rodriguez-Santiago, B., Casado, J. A., Rio, P., Bauser, C., Reina-Castillón, J., Lopez-Sanchez, M., Gonzalez-Quereda, L., Gallano, P., Catalá, A., Ruiz-Llobet, A., Badell, I., Diaz-Heredia, C., Hladun, R., Senent, L., Argiles, B., ... Surrallés, J. (2020). Optimised molecular genetic diagnostics of Fanconi anaemia by whole exome sequencing and functional studies. *Journal of Medical Genetics*, 57(4), 258–268. doi: [10.1136/jmedgenet-2019-106249](https://doi.org/10.1136/jmedgenet-2019-106249)
13. Bonfim, C. (2020). Special pre- and posttransplant considerations in inherited bone marrow failure and hematopoietic malignancy predisposition syndromes. *Hematology*, 2020(1), 107–114. doi: [10.1182/hematology.2020000095](https://doi.org/10.1182/hematology.2020000095)
14. Callén, E., Casado, J. A., Tischkowitz, M. D., Bueren, J. A., Creus, A., Marcos, R., Dasí, A., Estella, J. M., Muñoz, A., Ortega, J. J., de Winter, J., Joenje, H., Schindler, D., Hanenberg, H., Hodgson, S. V., Mathew, C. G., & Surrallés, J. (2005). A common founder mutation in FANCA underlies the world's highest prevalence of Fanconi anemia in Gypsy families from Spain. *Blood*, 105(5), 1946–1949. doi: [10.1182/blood-2004-07-2588](https://doi.org/10.1182/blood-2004-07-2588)
15. Camitta, B. M., Storb, R., & Thomas, E. D. (1982). Aplastic anemia: Pathogenesis, diagnosis, treatment, and Prognosis. *The New England Journal of Medicine*, 306(12), 712–718. doi: [10.1056/nejm198203253061204](https://doi.org/10.1056/nejm198203253061204)
16. Castella, M., Pujol, R., Callén, E., Trujillo, J. P., Casado, J. A., Gille, H., Lach, F. P., Auerbach, A. D., Schindler, D., Benítez, J., Porto, B., Ferro, T., Muñoz, A., Sevilla, J., Madero, L., Cela, E., Beléndez, C., de Heredia, C. D., Olivé, T., ... Surrallés, J. (2011). Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. *Blood*, 117(14), 3759–3769. doi: [10.1182/blood-2010-08-299917](https://doi.org/10.1182/blood-2010-08-299917)
17. Ceccaldi, R., Parmar, K., Mouly, E., Delord, M., Kim, J. M., Regairaz, M., Pla, M., Vasquez, N., Zhang, Q.-S., Pondarre, C., Peffault de Latour, R., Gluckman, E., Cavazzana-Calvo, M., Leblanc, T., Larghero, J., Grompe, M., Socié, G., D'Andrea, A. D., & Soulier, J. (2012). Bone marrow failure in Fanconi anemia is triggered by an exacerbated p53/p21 DNA damage response that impairs hematopoietic stem and progenitor cells. *Cell Stem Cell*, 11(1), 36–49. doi: [10.1016/j.stem.2012.05.013](https://doi.org/10.1016/j.stem.2012.05.013)
18. Ceccaldi, R., Sarangi, P., & D'Andrea, A. D. (2016). The Fanconi anaemia pathway: new players and new functions. *Nature Reviews. Molecular Cell Biology*, 17(6), 337–349. doi: [10.1038/nrm.2016.48](https://doi.org/10.1038/nrm.2016.48)
19. Chaudhury, I., Sareen, A., Raghunandan, M., & Sobeck, A. (2013). FANCD2 regulates BLM complex functions independently of FANCI to promote replication fork recovery. *Nucleic Acids Research*, 41(13), 6444–6459. doi: [10.1093/nar/gkt348](https://doi.org/10.1093/nar/gkt348)

20. Che, R., Zhang, J., Nepal, M., Han, B., & Fei, P. (2018). Multifaceted fanconi anemia signaling. *Trends in Genetics: TIG*, 34(3), 171–183. doi: [10.1016/j.tig.2017.11.006](https://doi.org/10.1016/j.tig.2017.11.006)
21. Cioc, A. M., Wagner, J. E., MacMillan, M. L., DeFor, T., & Hirsch, B. (2010). Diagnosis of myelodysplastic syndrome among a cohort of 119 patients with fanconi anemia: morphologic and cytogenetic characteristics: Morphologic and cytogenetic characteristics. *American Journal of Clinical Pathology*, 133(1), 92–100. doi: [10.1309/AJCP7W9VMJENZOVG](https://doi.org/10.1309/AJCP7W9VMJENZOVG)
22. de Winter, J. P., Léveillé, 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. *The American Journal of Human Genetics*, 67(5), 1306–1308. doi: [10.1016/S0002-9297\(07\)62959-0](https://doi.org/10.1016/S0002-9297(07)62959-0)
23. Winter, J. P., Rooimans, M. A., van Der Weel, L., van Berkel, C. G., Alon, N., Bosnoyan-Collins, L., de Groot, J., Zhi, Y., Waisfisz, Q., Pronk, J. C., Arwert, F., Mathew, C. G., Scheper, R. J., Hoatlin, M. E., Buchwald, M., & Joenje, H. (2000). The Fanconi anaemia gene FANCF encodes a novel protein with homology to ROM. *Nature Genetics*, 24(1), 15–16. doi: [10.1038/71626](https://doi.org/10.1038/71626)
24. 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. *Nature Genetics*, 20(3), 281–283. doi: [10.1038/3093](https://doi.org/10.1038/3093)
25. Deans, A. J., & West, S. C. (2011). DNA interstrand crosslink repair and cancer. *Nature Reviews. Cancer*, 11(7), 467–480. doi: [10.1038/nrc3088](https://doi.org/10.1038/nrc3088)
26. den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., Roux, A.-F., Smith, T., Antonarakis, S. E., & Taschner, P. E. M. (2016). HGVS recommendations for the description of sequence variants: 2016 update. *Human Mutation*, 37(6), 564–569. doi: [10.1002/humu.22981](https://doi.org/10.1002/humu.22981)
27. DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5), 491–498. doi: [10.1038/ng.806](https://doi.org/10.1038/ng.806)
28. DeZern, A. E., & Brodsky, R. A. (2011). Clinical management of aplastic anemia. *Expert Review of Hematology*, 4(2), 221–230. doi: [10.1586/ehm.11.11](https://doi.org/10.1586/ehm.11.11)
29. Dong, H., Nebert, D. W., Bruford, E. A., Thompson, D. C., Joenje, H., & Vasiliou, V. (2015). Update of the human and mouse Fanconi anemia genes. *Human Genomics*, 9(1). doi: [10.1186/s40246-015-0054-y](https://doi.org/10.1186/s40246-015-0054-y)

30. Donovan, F. X., Solanki, A., Mori, M., Chavan, N., George, M., C, S. K., Okuno, Y., Muramastu, H., Yoshida, K., Shimamoto, A., Takaori-Kondo, A., Yabe, H., Ogawa, S., Kojima, S., Yabe, M., Ramanagoudr-Bhojappa, R., Smogorzewska, A., Mohan, S., Rajendran, A., ... Vundinti, B. R. (2020). A founder variant in the South Asian population leads to a high prevalence of *FANCL* Fanconi anemia cases in India. *Human Mutation*, 41(1), 122–128. doi: [10.1002/humu.23914](https://doi.org/10.1002/humu.23914)
31. Du, W., Erden, O., & Pang, Q. (2014). TNF- α signaling in Fanconi anemia. *Blood Cells, Molecules & Diseases*, 52(1), 2–11. doi: [10.1016/j.bcmd.2013.06.005](https://doi.org/10.1016/j.bcmd.2013.06.005)
32. Dufour, C. (2017). How I manage patients with Fanconi anaemia. *British Journal of Haematology*, 178(1), 32–47. doi: [10.1111/bjh.14615](https://doi.org/10.1111/bjh.14615)
33. Dufour, C., & Pierri, F. (2022). Modern management of Fanconi anemia. *Hematology*, 2022(1), 649–657. doi: [10.1182/hematology.2022000393](https://doi.org/10.1182/hematology.2022000393)
34. Faivre, L., Guardiola, P., Lewis, C., Dokal, I., Ebell, W., Zatterale, A., Altay, C., Poole, J., Stones, D., Kwee, M. L., van Weel-Sipman, M., Havenga, C., Morgan, N., de Winter, J., Digweed, M., Savoia, A., Pronk, J., de Ravel, T., Jansen, S., ... Mathew, C. G. (2000). Association of complementation group and mutation type with clinical outcome in Fanconi anemia. European Fanconi Anemia Research Group. *Blood*, 96(13), 4064–4070. <https://pubmed.ncbi.nlm.nih.gov/11110674/>
35. Faivre, Laurence, Portnoï, M. F., Pals, G., Stoppa-Lyonnet, D., Le Merrer, M., Thauvin-Robinet, C., Huet, F., Mathew, C. G., Joenje, H., Verloes, A., & Baumann, C. (2005). Should chromosome breakage studies be performed in patients with VACTERL association?. *American Journal of Medical Genetics. Part A*, 137(1), 55–58. doi: [10.1002/ajmg.a.30853](https://doi.org/10.1002/ajmg.a.30853)
36. Fanconi anaemia/Breast cancer consortium. (1996). Positional cloning of the Fanconi anaemia group A gene. *Nature Genetics*, 14(3), 324–328. doi: [10.1038/ng1196-324](https://doi.org/10.1038/ng1196-324)
37. Fanconi, G (1927). Familiare, infantile, perniziosartige Anämie (pernizioses Blutbild und Konstitution). *Jahrbuch für Kinderheilkunde*; 117:257–80.
38. Feben, C., Haw, T., Stones, D., Jacobs, C., Sutton, C., Kromberg, J., & Krause, A. (2017). Fanconi anaemia in south African patients with Afrikaner ancestry. *SAJCH: The South African Journal of Child Health*, 11(3), 141. doi: [10.7196/sajch.2017.v11i3.1312](https://doi.org/10.7196/sajch.2017.v11i3.1312)
39. Fiesco-Roa, M. Ó., García-de Teresa, B., Leal-Anaya, P., van 't Hek, R., Wegman-Ostrosky, T., Frías, S., & Rodríguez, A. (2022). Fanconi anemia and dyskeratosis congenita/telomere biology disorders: Two inherited bone marrow failure syndromes with genomic instability. *Frontiers in Oncology*, 12, 949435. doi: [10.3389/fonc.2022.949435](https://doi.org/10.3389/fonc.2022.949435)
40. Fiesco-Roa, M. O., Giri, N., McReynolds, L. J., Best, A. F., & Alter, B. P. (2019). Genotype-phenotype associations in Fanconi anemia: A literature review. *Blood Reviews*, 37(100589), 100589. doi: [10.1016/j.blre.2019.100589](https://doi.org/10.1016/j.blre.2019.100589)

41. Flynn, E. K., Kamat, A., Lach, F. P., Donovan, F. X., Kimble, D. C., Narisu, N., Sanborn, E., Boulad, F., Davies, S. M., Gillio, A. P., 3rd, Harris, R. E., MacMillan, M. L., Wagner, J. E., Smogorzewska, A., Auerbach, A. D., Ostrander, E. A., & Chandrasekharappa, S. C. (2014). Comprehensive analysis of pathogenic deletion variants in Fanconi anemia genes. *Human Mutation*, 35(11), 1342–1353. [doi: 10.1002/humu.22680](https://doi.org/10.1002/humu.22680)
42. Fu, X., Tan, W., Song, Q., Pei, H., & Li, J. (2022). BRCA1 and breast cancer: Molecular mechanisms and therapeutic strategies. *Frontiers in Cell and Developmental Biology*, 10, 813457. [doi: 10.3389/fcell.2022.813457](https://doi.org/10.3389/fcell.2022.813457)
43. Futaki, M., Yamashita, T., Yagasaki, H., Toda, T., Yabe, M., Kato, S., Asano, S., & Nakahata, T. (2000). The IVS4 + 4 A to T mutation of the fanconi anemia gene *FANCC* is not associated with a severe phenotype in Japanese patients. *Blood*, 95(4), 1493–1498. <https://pubmed.ncbi.nlm.nih.gov/10666230/>
44. Garaycochea, J. I., & Patel, K. J. (2014). Why does the bone marrow fail in Fanconi anemia? *Blood*, 123(1), 26–34. [doi: 10.1182/blood-2013-09-427740](https://doi.org/10.1182/blood-2013-09-427740)
45. Garaycochea, J. I., Crossan, G. P., Langevin, F., Mulderrig, L., Louzada, S., Yang, F., Guilbaud, G., Park, N., Roerink, S., Nik-Zainal, S., Stratton, M. R., & Patel, K. J. (2018). Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature*, 553(7687), 171–177. [doi: 10.1038/nature25154](https://doi.org/10.1038/nature25154)
46. García de Teresa, B., Rodríguez, A., & Frías, S. (2016). Estudio multidisciplinario del paciente con anemia de Fanconi. *Acta Pediátrica de México*, 37(1), 54. [doi: 10.18233/apm37no1pp54-59](https://doi.org/10.18233/apm37no1pp54-59)
47. García-de-Teresa, B., Rodríguez, A., & Frías, S. (2020). Chromosome instability in fanconi anemia: From breaks to phenotypic consequences. *Genes*, 11(12), 1528. [doi: 10.3390/genes11121528](https://doi.org/10.3390/genes11121528)
48. Garza-Mayén, G (2021): Presencia de daño en el DNA en células troncales y progenitoras hematopoyéticas de pacientes con Anemia de Fanconi y su correlación con la progresión de falla medular (tesis de maestría inédita). Universidad Anahuac (México).
49. Gordon, S. M., & Buchwald, M. (2013). *The FANCC gene and its products*. Philadelphia, PA: Landes Bioscience.
50. Gulbis, B., Eleftheriou, A., Angastiniotis, M., Ball, S., Surrallés, J., Castella, M., Heimpel, H., Hill, A., & Corrons, J.-L. V. (2010). Epidemiology of rare anaemias in Europe. *Advances in Experimental Medicine and Biology*, 686, 375–396. [doi: 10.1007/978-90-481-9485-8_22](https://doi.org/10.1007/978-90-481-9485-8_22)
51. Helbling-Leclerc, A., Garcin, C., & Rosselli, F. (2021). Beyond DNA repair and chromosome instability-Fanconi anaemia as a cellular senescence-associated syndrome. *Cell Death and Differentiation*, 28(4), 1159–1173. [doi: 10.1038/s41418-021-00764-5](https://doi.org/10.1038/s41418-021-00764-5)

52. Hu, J., & Ng, P. C. (2012). Predicting the effects of frameshifting indels. *Genome Biology*, 13(2), R9. <https://doi.org/10.1186/gb-2012-13-2-r9>
53. Joshi, G., Arthur, N. B. J., Geetha, T. S., Datari, P. V. R., Modak, K., Roy, D., Chaudhury, A. D., Sundaraganesan, P., Priyanka, S., Na, F., Ramprasad, V., Abraham, A., Srivastava, V. M., Srivastava, A., Kulkarni, U. P., George, B., & Velayudhan, S. R. (2023). Comprehensive laboratory diagnosis of Fanconi anaemia: comparison of cellular and molecular analysis. *Journal of Medical Genetics*, 60(8), 801–809. <https://doi.org/10.1136/jmg-2022-108714>
54. Jung, M., Ramanagoudr-Bhojappa, R., van Twest, S., Rosti, R. O., Murphy, V., Tan, W., Donovan, F. X., Lach, F. P., Kimble, D. C., Jiang, C. S., Vaughan, R., Mehta, P. A., Pierri, F., Dufour, C., Auerbach, A. D., Deans, A. J., Smogorzewska, A., & Chandrasekharappa, S. C. (2020). Association of clinical severity with FANCB variant type in Fanconi anemia. *Blood*, 135(18), 1588–1602. [doi: 10.1182/blood.2019003249](https://doi.org/10.1182/blood.2019003249)
55. Katsonis, P., Wilhelm, K., Williams, A., & Lichtarge, O. (2022). Genome interpretation using in silico predictors of variant impact. *Human Genetics*, 141(10), 1549–1577. <https://doi.org/10.1007/s00439-022-02457-6>
56. Kelly, P. F., Radtke, S., von Kalle, C., Balcik, B., Bohn, K., Mueller, R., Schuesler, T., Haren, M., Reeves, L., Cancelas, J. A., Leemhuis, T., Harris, R., Auerbach, A. D., Smith, F. O., Davies, S. M., & Williams, D. A. (2007). Stem cell collection and gene transfer in Fanconi anemia. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, 15(1), 211–219. [doi: 10.1038/sj.mt.6300033](https://doi.org/10.1038/sj.mt.6300033)
57. Kennedy, R. D., & D'Andrea, A. D. (2005). The Fanconi Anemia/BRCA pathway: new faces in the crowd. *Genes & Development*, 19(24), 2925–2940. [doi: 10.1101/gad.1370505](https://doi.org/10.1101/gad.1370505)
58. Kimble, D. C., Lach, F. P., Gregg, S. Q., Donovan, F. X., Flynn, E. K., Kamat, A., Young, A., Vemulapalli, M., Thomas, J. W., Mullikin, J. C., Auerbach, A. D., Smogorzewska, A., & Chandrasekharappa, S. C. (2018). A comprehensive approach to identification of pathogenic FANCA variants in Fanconi anemia patients and their families. *Human Mutation*, 39(2), 237–254. [doi: 10.1002/humu.23366](https://doi.org/10.1002/humu.23366)
59. Kumari, U., Ya Jun, W., Huat Bay, B., & Lyakhovich, A. (2014). Evidence of mitochondrial dysfunction and impaired ROS detoxifying machinery in Fanconi anemia cells. *Oncogene*, 33(2), 165–172. [doi: 10.1038/onc.2012.583](https://doi.org/10.1038/onc.2012.583)
60. Kutler, D. I., & Auerbach, A. D. (2004). Fanconi anemia in Ashkenazi Jews. *Familial Cancer*, 3(3–4), 241–248. [doi: 10.1007/s10689-004-9565-8](https://doi.org/10.1007/s10689-004-9565-8)
61. Kutler, D. I., Auerbach, A. D., Satagopan, J., Giampietro, P. F., Batish, S. D., Huvos, A. G., Goberdhan, A., Shah, J. P., & Singh, B. (2003). High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. *Archives of*

- Otolaryngology--Head & Neck Surgery*, 129(1), 106–112. doi: [10.1001/archotol.129.1.106](https://doi.org/10.1001/archotol.129.1.106)
62. 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(4), 1249–1256. doi: [10.1182/blood-2002-07-2170](https://doi.org/10.1182/blood-2002-07-2170)
 63. Lee, A.-Y. (2021). Skin pigmentation abnormalities and their possible relationship with skin aging. *International Journal of Molecular Sciences*, 22(7), 3727. doi: [10.3390/ijms22073727](https://doi.org/10.3390/ijms22073727)
 64. Li, X., Wilson, A. F., Du, W., & Pang, Q. (2018). Cell-cycle-specific function of p53 in Fanconi anemia hematopoietic stem and progenitor cell proliferation. *Stem Cell Reports*, 10(2), 339–346. doi: [10.1016/j.stemcr.2017.12.006](https://doi.org/10.1016/j.stemcr.2017.12.006)
 65. Mamrak, N. E., Shimamura, A., & Howlett, N. G. (2017). Recent discoveries in the molecular pathogenesis of the inherited bone marrow failure syndrome Fanconi anemia. *Blood Reviews*, 31(3), 93–99. doi: [10.1016/j.blre.2016.10.002](https://doi.org/10.1016/j.blre.2016.10.002)
 66. Matsui, W. H., Brodsky, R. A., Smith, B. D., Borowitz, M. J., & Jones, R. J. (2006). Quantitative analysis of bone marrow CD34 cells in aplastic anemia and hypoplastic myelodysplastic syndromes. *Leukemia*, 20(3), 458–462. doi: [10.1038/sj.leu.2404119](https://doi.org/10.1038/sj.leu.2404119)
 67. Meetei, A. R., Levitus, M., Xue, Y., Medhurst, A. L., Zwaan, M., Ling, C., Roomans, 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. *Nature Genetics*, 36(11), 1219–1224. doi: [10.1038/ng1458](https://doi.org/10.1038/ng1458)
 68. Juárez-Figueroa, U; Ayala-Zambrano, C; Reyes, P; Frías, S. (2018). Origin and consequences of genomic instability. Chromosomal instability syndromes. *Mensaje Bioquímico*, 42, 64-80. <http://bq.facmed.unam.mx/tab/wp-content/uploads/2020/02/7-Dra.-Sara-Frias.pdf>
 69. McReynolds, L. J., Biswas, K., Giri, N., Sharan, S. K., & Alter, B. P. (2021). Genotype-cancer association in patients with Fanconi anemia due to pathogenic variants in FANCD1 (BRCA2) or FANCN (PALB2). *Cancer Genetics*, 258–259, 101–109. <https://doi.org/10.1016/j.cancergen.2021.10.001>
 70. Merfort, L. W., Lisboa, M. de O., Cavalli, L. R., & Bonfim, C. M. S. (2022). Cytogenetics in fanconi Anemia: The importance of follow-up and the search for new biomarkers of genomic instability. *International Journal of Molecular Sciences*, 23(22), 14119. doi: [10.3390/ijms232214119](https://doi.org/10.3390/ijms232214119)
 71. Milletti, G., Strocchio, L., Pagliara, D., Girardi, K., Carta, R., Mastronuzzi, A., Locatelli, F., & Nazio, F. (2020). Canonical and noncanonical roles of fanconi anemia proteins: Implications in cancer predisposition. *Cancers*, 12(9), 2684. doi: [10.3390/cancers12092684](https://doi.org/10.3390/cancers12092684)
 72. Molina, B., Frías, S., & Ramos, S. (2022). Anemia de Fanconi, Parte 1. Diagnóstico citogenético. *Acta Pediátrica de México*, 43(2), 102. doi: [10.18233/apm43no2pp102-1282393](https://doi.org/10.18233/apm43no2pp102-1282393)

73. Morgan, N. V., Essop, F., Demuth, I., de Ravel, T., Jansen, S., Tischkowitz, M., Lewis, C. M., Wainwright, L., Poole, J., Joenje, H., Digweed, M., Krause, A., & Mathew, C. G. (2005). A common Fanconi anemia mutation in black populations of sub-Saharan Africa. *Blood*, 105(9), 3542–3544. doi: [10.1182/blood-2004-10-3968](https://doi.org/10.1182/blood-2004-10-3968)
74. Nalepa, G., & Clapp, D. W. (2018). Fanconi anaemia and cancer: an intricate relationship. *Nature Reviews in Cancer*, 18(3), 168–185. doi: [10.1038/nrc.2017.116](https://doi.org/10.1038/nrc.2017.116)
75. Nalepa, G., Enzor, R., Sun, Z., Marchal, C., Park, S.-J., Yang, Y., Tedeschi, L., Kelich, S., Hanenberg, H., & Clapp, D. W. (2013). Fanconi anemia signaling network regulates the spindle assembly checkpoint. *The Journal of Clinical Investigation*, 123(9), 3839–3847. doi: [10.1172/JCI67364](https://doi.org/10.1172/JCI67364)
76. Neveling, K., Endt, D., Hoehn, H., & Schindler, D. (2009). Genotype-phenotype correlations in Fanconi anemia. *Mutation Research*, 668(1–2), 73–91. doi: [10.1016/j.mrfmmm.2009.05.006](https://doi.org/10.1016/j.mrfmmm.2009.05.006)
77. Nie, D., Zhang, J., Wang, F., Zhang, W., Liu, L., Chen, X., Zhang, Y., Cao, P., Xiong, M., Wang, T., Wu, P., Ma, X., Tian, W., Wang, M., Chen, K. N., & Liu, H. (2020). Comprehensive analysis on phenotype and genetic basis of Chinese Fanconi anemia patients: dismal outcomes call for nationwide studies. *BMC Medical Genetics*, 21(1). <https://doi.org/10.1186/s12881-020-01057-3>
78. Niraj, J., Färkkilä, A., & D’Andrea, A. D. (2019). The fanconi anemia pathway in cancer. *Annual Review of Cancer Biology*, 3(1), 457–478. doi: [10.1146/annurev-cancerbio-030617-050422](https://doi.org/10.1146/annurev-cancerbio-030617-050422)
79. Oostra, A. B., Nieuwint, A. W. M., Joenje, H., & de Winter, J. P. (2012). Diagnosis of fanconi anemia: chromosomal breakage analysis. *Anemia*, 2012, 238731. doi: [10.1155/2012/238731](https://doi.org/10.1155/2012/238731)
80. Pagano, G., Tiano, L., Pallardó, F. V., Lyakhovich, A., Mukhopadhyay, S. S., Di Bartolomeo, P., Zatterale, A., & Trifuoggi, M. (2021). Re-definition and supporting evidence toward Fanconi Anemia as a mitochondrial disease: Prospects for new design in clinical management. *Redox Biology*, 40(101860), 101860. doi: [10.1016/j.redox.2021.101860](https://doi.org/10.1016/j.redox.2021.101860)
81. Peake, J. D., & Noguchi, E. (2022). Fanconi anemia: current insights regarding epidemiology, cancer, and DNA repair. *Human Genetics*, 141(12), 1811–1836. doi: [10.1007/s00439-022-02462-9](https://doi.org/10.1007/s00439-022-02462-9)
82. Peake, J. D., & Noguchi, E. (2022). Fanconi anemia: current insights regarding epidemiology, cancer, and DNA repair. *Human Genetics*, 141(12), 1811–1836. doi: [10.1007/s00439-022-02462-9](https://doi.org/10.1007/s00439-022-02462-9)
83. Peffault de Latour, R., & Soulier, J. (2016). How I treat MDS and AML in Fanconi anemia. *Blood*, 127(24), 2971–2979. doi: [10.1182/blood-2016-01-583625](https://doi.org/10.1182/blood-2016-01-583625)
84. Pillionetto, D. V., Piovezan, B. Z., Nichele, S., Lima, A. C. M., Pasquini, R., Pereira, N. F., & Bonfim, C. (2022). Investigation of mutations in Fanconi anemia genes and malignancy predisposition in Brazilian patients. *International Journal of Laboratory Hematology*. doi: [10.1111/ijlh.13986](https://doi.org/10.1111/ijlh.13986)

85. 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. *Nature Genetics*, 39(2), 162–164. doi: [10.1038/ng1947](https://doi.org/10.1038/ng1947)
86. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H. L., & ACMG Laboratory Quality Assurance Committee. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 17(5), 405–424. doi: [10.1038/gim.2015.30](https://doi.org/10.1038/gim.2015.30)
87. Riolo, G., Cantara, S., & Ricci, C. (2021). What’s wrong in a jump? Prediction and validation of splice site variants. *Methods and Protocols*, 4(3), 62. <https://doi.org/10.3390/mps4030062>
88. Rodríguez, A., & D’Andrea, A. (2017). Fanconi anemia pathway. *Current Biology: CB*, 27(18), R986–R988. doi: [10.1016/j.cub.2017.07.043](https://doi.org/10.1016/j.cub.2017.07.043)
89. Rodríguez, A., Epperly, M., Filiatrault, J., Velázquez, M., Yang, C., McQueen, K., Sambel, L. A., Nguyen, H., Iyer, D. R., Juárez, U., Ayala-Zambrano, C., Martignetti, D. B., Frías, S., Fisher, R., Parmar, K., Greenberger, J. S., & D’Andrea, A. D. (2022). TGFβ pathway is required for viable gestation of Fanconi anemia embryos. *PLoS Genetics*, 18(11), e1010459. doi: [10.1371/journal.pgen.1010459](https://doi.org/10.1371/journal.pgen.1010459)
90. Rodríguez, A., Zhang, K., Färkkilä, A., Filiatrault, J., Yang, C., Velázquez, M., Furutani, E., Goldman, D. C., García de Teresa, B., Garza-Mayén, G., McQueen, K., Sambel, L. A., Molina, B., Torres, L., González, M., Vadillo, E., Pelayo, R., Fleming, W. H., Grompe, M., ... D’Andrea, A. D. (2021). MYC promotes bone marrow stem cell dysfunction in Fanconi anemia. *Cell Stem Cell*, 28(1), 33-47.e8. doi: [10.1016/j.stem.2020.09.004](https://doi.org/10.1016/j.stem.2020.09.004)
91. Rosenberg, P. S., Greene, M. H., & Alter, B. P. (2003). Cancer incidence in persons with Fanconi anemia. *Blood*, 101(3), 822–826. doi: [10.1182/blood-2002-05-1498](https://doi.org/10.1182/blood-2002-05-1498)
92. Rosenberg, P. S., Tamary, H., & Alter, B. P. (2011). How high are carrier frequencies of rare recessive syndromes? Contemporary estimates for Fanconi Anemia in the United States and Israel. *American Journal of Medical Genetics. Part A*, 155A(8), 1877–1883. doi: [10.1002/ajmg.a.34087](https://doi.org/10.1002/ajmg.a.34087)
93. Roy, S., Coldren, C., Karunamurthy, A., Kip, N. S., Klee, E. W., Lincoln, S. E., Leon, A., Pullambhatla, M., Temple-Smolkin, R. L., Voelkerding, K. V., Wang, C., & Carter, A. B. (2018). Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: A joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. *The Journal of Molecular Diagnostics: JMD*, 20(1), 4–27. doi: [10.1016/j.jmoldx.2017.11.003](https://doi.org/10.1016/j.jmoldx.2017.11.003)

94. Savage, S. A., & Walsh, M. F. (2018). Myelodysplastic syndrome, acute myeloid leukemia, and cancer surveillance in fanconi anemia. *Hematology/Oncology Clinics of North America*, 32(4), 657–668. [doi: 10.1016/j.hoc.2018.04.002](https://doi.org/10.1016/j.hoc.2018.04.002)
95. Schroeder, T (1966). Cytogenic and cytology in enzymopenic panmyelopathies and pancitopenias. Familial myelopathy of Fanconi, glutathione-reductase deficiency anemia and megaloblastic B12 deficiency anemia. *Humangenetik* 2(3):287-316. [doi:10.1007/bf00395947](https://doi.org/10.1007/bf00395947)
96. Schuler, D; Kiss, A y Fabian, A (1969). Chromosome studies in Fanconi anemia. *Orv Hetil* 110(13):713-20 passim. <https://pubmed.ncbi.nlm.nih.gov/5768668/>
97. Shakeel, S., Rajendra, E., Alcón, P., O'Reilly, F., Chorev, D. S., Maslen, S., Degliesposti, G., Russo, C. J., He, S., Hill, C. H., Skehel, J. M., Scheres, S. H. W., Patel, K. J., Rappsilber, J., Robinson, C. V., & Passmore, L. A. (2019). Structure of the Fanconi anaemia monoubiquitin ligase complex. *Nature*, 575(7781), 234–237. [doi: 10.1038/s41586-019-1703-4](https://doi.org/10.1038/s41586-019-1703-4)
98. Shukla, P., Korgaonkar, S., Kerketta, L., & Vundinti, B. R. (2018). Prevalence and association of dermatological manifestations with fanconi anemia: A retrospective study. *Indian Dermatology Online Journal*, 9(5), 341–342. [doi: 10.4103/idoj.IDOJ_368_17](https://doi.org/10.4103/idoj.IDOJ_368_17)
99. Solanki, A., Kumar Selvaa, C., Sheth, F., Radhakrishnan, N., Kalra, M., & Vundinti, B. R. (2017). Characterization of two novel FANCG mutations in Indian Fanconi anemia patients. *Leukemia Research*, 53, 50–56. [doi: 10.1016/j.leukres.2016.11.013](https://doi.org/10.1016/j.leukres.2016.11.013)
100. Strathdee, C. A., Duncan, A. M., & Buchwald, M. (1992). Evidence for at least four Fanconi anaemia genes including FACC on chromosome 9. *Nature Genetics*, 1(3), 196–198. [doi: 10.1038/ng0692-196](https://doi.org/10.1038/ng0692-196)
101. Steinberg-Shemer, O., Goldberg, T. A., Yacobovich, J., Levin, C., Koren, A., Revel-Vilk, S., Ben-Ami, T., Kuperman, A. A., Zemer, V. S., Toren, A., Kapelushnik, J., Ben-Barak, A., Miskin, H., Krasnov, T., Noy-Lotan, S., Dgany, O., & Tamary, H. (2020). Characterization and genotype-phenotype correlation of patients with Fanconi anemia in a multi-ethnic population. *Haematologica*, 105(7), 1825–1834. <https://doi.org/10.3324/haematol.2019.222877>
102. Swuec, P., Renault, L., Borg, A., Shah, F., Murphy, V. J., van Twest, S., Snijders, A. P., Deans, A. J., & Costa, A. (2017). The FA core complex contains a homo-dimeric catalytic module for the symmetric mono-ubiquitination of FANCI-FANCD2. *Cell Reports*, 18(3), 611–623. [doi: 10.1016/j.celrep.2016.11.013](https://doi.org/10.1016/j.celrep.2016.11.013)
103. Tan, W., van Twest, S., Murphy, V. J., & Deans, A. J. (2020). ATR-mediated FANCI phosphorylation regulates both ubiquitination and deubiquitination of FANCD2. *Frontiers in Cell and Developmental Biology*, 8, 2. [doi: 10.3389/fcell.2020.00002](https://doi.org/10.3389/fcell.2020.00002)
104. 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. *Molecular Cell*, 7(2), 241–248. [doi: 10.1016/s1097-2765\(01\)00172-1](https://doi.org/10.1016/s1097-2765(01)00172-1)

105. 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. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10), 5734–5739. [doi: 10.1073/pnas.091402398](https://doi.org/10.1073/pnas.091402398)
106. Tischkowitz, M. D., Morgan, N. V., Grimwade, D., Eddy, C., Ball, S., Vorechovsky, I., Langabeer, S., Stöger, R., Hodgson, S. V., & Mathew, C. G. (2004). Deletion and reduced expression of the Fanconi anemia FANCA gene in sporadic acute myeloid leukemia. *Leukemia*, 18(3), 420–425. <https://doi.org/10.1038/sj.leu.2403280>
107. Trottier, M., & Meyn, S. (2008). Fanconi anemia and the response of human hematopoietic cells to DNA damage. *Blood*, 112(11), 441–441. [doi: 10.1182/blood.v112.11.441.441](https://doi.org/10.1182/blood.v112.11.441.441)
108. Tryon, R., Zierhut, H., MacMillan, M. L., & Wagner, J. E. (2017). Phenotypic variability in patients with Fanconi anemia and biallelic FANCF mutations. *American Journal of Medical Genetics. Part A*, 173(1), 260–263. <https://doi.org/10.1002/ajmg.a.37998>
109. Tsui, V., & Crismani, W. (2019). The fanconi anemia pathway and fertility. *Trends in Genetics: TIG*, 35(3), 199–214. [doi: 10.1016/j.tig.2018.12.007](https://doi.org/10.1016/j.tig.2018.12.007)
110. Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, 43(1), 11.10.1-11.10.33. [doi: 10.1002/0471250953.bi1110s43](https://doi.org/10.1002/0471250953.bi1110s43)
111. van Twest, S., Murphy, V. J., Hodson, C., Tan, W., Swuec, P., O'Rourke, J. J., Heierhorst, J., Crismani, W., & Deans, A. J. (2017). Mechanism of ubiquitination and deubiquitination in the fanconi anemia pathway. *Molecular Cell*, 65(2), 247–259. [doi: 10.1016/j.molcel.2016.11.005](https://doi.org/10.1016/j.molcel.2016.11.005)
112. Wajnrajch, M. P., Gertner, J. M., Huma, Z., Popovic, J., Lin, K., Verlander, P. C., Batish, S. D., Giampietro, P. F., Davis, J. G., New, M. I., & Auerbach, A. D. (2001). Evaluation of growth and hormonal status in patients referred to the International Fanconi Anemia Registry. *Pediatrics*, 107(4), 744–754. [doi: 10.1542/peds.107.4.744](https://doi.org/10.1542/peds.107.4.744)
113. 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. *Nature Genetics*, 4(2), 202–205. [doi: 10.1038/ng0693-202](https://doi.org/10.1038/ng0693-202)
114. Woodward, E. R., & Meyer, S. (2021). Fanconi anaemia, childhood cancer and the BRCA genes. *Genes*, 12(10), 1520. [doi: 10.3390/genes12101520](https://doi.org/10.3390/genes12101520)
115. Zareifar, S., Dastsooz, H., Shahriari, M., Faghihi, M. A., Shekarkhar, G., Bordbar, M., Zekavat, O. R., & Shakibzad, N. (2019). A novel frame-shift

- deletion in *FANCF* gene causing autosomal recessive Fanconi anemia: a case report. *BMC Medical Genetics*, 20(1), 122. doi: [10.1186/s12881-019-0855-2](https://doi.org/10.1186/s12881-019-0855-2)
116. Zhang, J., Dewar, J. M., Budzowska, M., Motnenko, A., Cohn, M. A., & Walter, J. C. (2015). DNA interstrand cross-link repair requires replication-fork convergence. *Nature Structural & Molecular Biology*, 22(3), 242–247. doi: [10.1038/nsmb.2956](https://doi.org/10.1038/nsmb.2956)

14. ANEXOS.

14.1 Tablas suplementarias.

Tabla S1: Programas utilizados para evaluar la patogenicidad de las variantes.

Programa	Fuente (liga de acceso)	Principio
Variantes no sinónimas (de sentido erróneo).		
SIFT 4.0	https://sift.bii.a-star.edu.sg/	Comparación de secuencias; análisis bayesianos.
Polyphen 2.0	http://genetics.bwh.harvard.edu/pph2/	Clasificador bayesiano ingenuo; aprendizaje automático.
MutPred2	http://mutpred.mutdb.org/	Aprendizaje automático supervisado, Árboles predictores ("Random Forest").
SNAP2	https://rostlab.org/services/snap2web/	Inteligencia artificial; uso de redes neuronales
MuPro	http://mupro.proteomics.ics.uci.edu/	Aprendizaje automático supervisado, vectores de soporte.
I-Mutant 3.0	http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi	Aprendizaje automático supervisado; vectores de soporte.
Consurf	https://consurf.tau.ac.il/consurf_index.php	Alineamiento múltiple de secuencias
PHANTER	http://www.pantherdb.org/	Modelos ocultos de Márkov y árboles de decisión.
Variantes sin sentido y de corrimiento de marco de lectura		
VEST-indel	https://karchinlab.org/apps/appvest.html	Aprendizaje automático supervisado ; Árboles predictores ("Random Forest")
MudPredLOF	http://mutpred2.mutdb.org/mutpredlof/#qform	Inteligencia artificial mediante redes neuronales
Variantes que afectan el proceso de corte y empalme		
HSF	http://www.umd.be/HSF3/	Matriz de pesos posicionales y algoritmos de entropía.
MaxEntScan	http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html	Distribución máxima de entropía.
CRYP-SKYP	https://cryp-skip.img.cas.cz/	Predicción a través de regresión logística múltiple.
RNABPS	https://home.jbnu.ac.kr/NSCL/rnabps.htm	Algoritmos de aprendizaje profundo ("Deep learning")
Representación de interacciones en modelos de proteínas.		
PyMol	https://pymol.org/2/	Visor molecular de código abierto
AlphaFold	https://alphafold.ebi.ac.uk/	Inteligencia artificial; redes neuronales.
Dynamut	https://biosig.lab.uq.edu.au/dynamut/	Aprendizaje automático supervisado; dinámica de proteínas

Tabla S2. Evaluación de la patogenicidad de variantes de corrimiento de marco de lectura y de aquellas que conducen a codón de paro.

Gen	Número NP	Cambio	MutPredLOF	VEST-indel
<i>FANCA</i>	NP_000126.1	p.L61WfsTer4	0.4578	0.715
		p.C846QfsTer20	0.4356	0.821
		p.C80VfsTer15	0.4592	0.734
		p.E440X	0.5128	0.937
<i>FANCC</i>	NP_000127.2	p.D23lfsTer23	0.3557	0.598
<i>FANCF</i>	NP_073562.1	p.R29PfsTer52	0.379	0.345
<i>FANCB</i>	NP_001018123.1	p.W479X	0.3824	0.927
		p.F118lfsTer5	0.3284	0.846
<i>FANCD2</i>	NP_001018125.1	p.Q1053GfsTer31	0.3996	0.847
<i>FANCD1</i>	NP_000050.3	p.V220lfsTer4	0.419	0.724
	NP_000050.3	p.Q1089SfsTer10	0.441	0.877
<i>FANCI</i>	NP_114432.2	p.R798X	0.5205	0.903

En la tabla: NP: Código identificador de la proteína, de acuerdo a las categorías de datos sugeridas en la base de datos RefSeq. Los cambios en los aminoácidos se etiquetaron de acuerdo a la nomenclatura de una letra; fs: corrimiento en el marco de lectura; X: presencia de codón de paro. En el programa MutPredLOF, el valor mostrado corresponde al promedio de las estimaciones que utiliza la red neural para evaluar el impacto de la variante sobre las propiedades de la proteína. Un valor de 0.5 (umbral) sugiere patogenicidad. En el programa VEST-indel, que utiliza enfoques de aprendizaje automático (“machine learning”) supervisado, un valor superior a 0.5 indica patogenicidad.

S2. Evaluación de patogenicidad en variantes de corte y empalme.

Gen	Cambio	MaxEntScan	CRYP-SKYP	Efecto
FANCA	c.1126-2A>C	4.71	0.35	Exon Skipping
	c.4260+1G>A	9.65	0.66	Nuevo sitio críptico
	c.827-1G>A	5.4	0.42	Exon Skipping
FANCG	c.511-3_511-2delCA	9.63	0.23	Exon Skipping

En la tabla: El valor de entropía calculado por MaxEntScan (Max Entropy Scan) establece que mientras el estimador sea superior a 3.5, mayor es la disparidad entre la muestra de referencia y la de alteración, alterando el proceso de corte y empalme. El valor que se muestra para el programa CRYP-SKYP corresponde al estimador PCR-E (probabilidad de activación de un sitio críptico de corte y empalme). Un valor inferior a 0.5 indica aumento de probabilidad de salto del exón (*skipping*), mientras que un valor superior a 0.5 indica mayor probabilidad de un sitio críptico (cambio en el patrón normal de remoción de intrones).

Tabla S3. Determinación de haplotipo en pacientes con variante fundadora en FANCG.

Haplotype FANCG Mixe T C T C G T C T G A T A C G T C C T T T A G T G G C A T C G Haplotypes in MXL	Count	Frequency
T G C A T C T G T A T A T T T A C T T A A G T G C G G T A C	2	0.0156
C C C A T C T G T G T A C T T A T C T G G G C G C G G C G G	1	0.0078
C G C A T C T A C A C G T G A A C C G G G T T A C G G C G C	1	0.0078
C G C A T C T A C A C G T G A C T C T A G T C A C G A T G C	1	0.0078
C G C A T C T A C A C G T T A A T C G G G G T A C C G C A C	1	0.0078
C G C A T C T A C A T A T T A A T C G G A G T A C G G C A G	1	0.0078
C G C A T C T A C G C G C T A A C C G A G T T A C G G C A G	1	0.0078
C G C A T C T A C G C G C T A A T C T G A T T A C G A C A C	1	0.0078
C G C A T C T A C G C G C T A A T C T G A T T A T G G C A C	1	0.0078
C G C A T C T A C G C G C T A C T C T A G T C A C G A T G C	1	0.0078
C G C A T C T A C G C G T T A A T C G G A G T G C G A T G C	1	0.0078
C G C A T C T A C G C G T T A A T C G G G T T G C G G C G G	1	0.0078
C G C A T C T A C G T A C G A A T C G G G T T A C G G C A C	1	0.0078
C G C A T C T A C G T A T G A A T C T G G T C G C G A C G C	1	0.0078
C G C A T C T A C G T A T T A A T C G A A G T G C G A C G C	1	0.0078
C G C A T C T A C G T A T T A A T C G G G T C A T G A T A G	1	0.0078
C G C A T C T A C G T A T T A A T C T A G T C A C G G C A C	1	0.0078
C G C A T C T A C G T A T T A A T C T G A T C A C G G C A C	1	0.0078
C G C A T C T G T A C G T G A A T C G G G T T A C G G C A C	1	0.0078
C G C A T C T G T A C G T T A A T C T G A T T A C C G T G C	1	0.0078
C G C A T C T G T A C G T T A A T C T G G T T A C G G C A G	1	0.0078
C G C A T C T G T A T A T G A A C C G G A G T A C C G C A G	1	0.0078
C G C A T C T G T A T A T T A A C C G G A T T A C G A T A C	1	0.0078
C G C A T C T G T A T A T T A A C C G G G G C G C G A C A C	1	0.0078
C G C A T C T G T A T A T T A A T C G G G T C A C G A C A G	1	0.0078
C G C A T C T G T A T A T T A C T C G G G G C A C G G C G C	1	0.0078
C G C A T C T G T G C G C G A A T C G A G T T G C C G C A C	1	0.0078
C G C A T C T G T G C G C T A A C C G G A T C G C G G T A C	1	0.0078
C G C A T C T G T G C G C T A C T C G G A T C A C G A T G C	1	0.0078
C G C A T C T G T G C G T T A A C T T G A T T A C G G C G C	1	0.0078
C G C A T C T G T G C G T T A A T C G G A T T A C G G C A C	1	0.0078
C G C A T C T G T G C G T T A A T C G G G G C A C G G T A C	1	0.0078
C G C A T C T G T G C G T T A A T C G G G T C G C G G T G C	1	0.0078
C G C A T C T G T G T A C T A A C C T G G T C G C G G C A C	1	0.0078
C G C A T C T G T G T A C T A A T C T G G T C A C G A C A C	1	0.0078
C G C A T C T G T G T A T G A A C C G G G T C G C G A C A C	1	0.0078
C G C A T C T G T G T A T T A A C C T A G G T A C C G C G C	1	0.0078
C G C A T C T G T G T A T T A A T C G G G G C G C G G C A C	1	0.0078
C G C A T C T G T G T A T T A A T C G G G G T A C C G C G G	1	0.0078
C G C A T C T G T G T A T T A A T T G G A G T A C G G C A C	1	0.0078
C G C A T T C A C A C G T T A A T C G G G T C A T G A T G C	1	0.0078
C G C A T T C A C A T A T T A A C C G G G G C A C G G T A C	1	0.0078
C G C A T T C A C G C G T T T A C T G A A T C A C G G T G C	1	0.0078

C_G_C_A_T_T_C_A_C_G_T_A_T_T_A_A_T_C_T_A_G_T_C_A_C_G_G_T_A_C	1	0.0078
C_G_C_A_T_T_C_G_C_A_C_G_T_G_A_A_T_C_T_G_G_G_C_A_C_G_G_C_A_C	1	0.0078
C_G_C_A_T_T_C_G_T_A_C_G_C_T_A_A_T_C_G_G_G_T_C_G_C_G_G_T_A_C	1	0.0078
C_G_C_A_T_T_C_G_T_A_C_G_C_T_A_C_T_C_G_G_G_G_C_A_C_G_A_T_G_C	1	0.0078
C_G_C_A_T_T_C_G_T_A_C_G_C_T_A_C_T_C_G_G_G_T_C_A_C_G_G_T_A_C	1	0.0078
C_G_C_A_T_T_C_G_T_A_C_G_T_T_A_C_C_C_G_G_G_T_C_G_C_G_A_C_A_C	1	0.0078
C_G_C_A_T_T_C_G_T_A_T_A_C_T_A_A_C_C_G_G_G_G_C_A_C_G_A_C_G_G	1	0.0078
C_G_T_G_G_C_T_A_C_A_C_G_T_T_A_C_C_C_G_G_G_T_C_G_C_G_A_T_A_C	1	0.0078
C_G_T_G_G_C_T_A_C_A_C_G_T_T_A_C_C_C_G_G_G_T_C_G_C_G_A_T_A_C	1	0.0078
C_G_T_G_G_C_T_A_C_G_C_G_C_T_A_A_C_C_G_A_G_G_C_A_C_G_G_C_A_C	1	0.0078
C_G_T_G_G_C_T_A_C_G_C_G_T_G_A_A_C_C_G_G_G_T_G_C_G_G_C_G_C	1	0.0078
C_G_T_G_G_C_T_A_C_G_C_G_T_G_A_A_C_C_G_G_G_T_G_C_G_G_C_G_C	1	0.0078

14.2 Artículos publicados.

- A. Juárez-Figueroa, U; Ayala-Zambrano, C; Reyes, P; Frías, S. (2018). Origin and consequences of genomic instability. Chromosomal instability syndromes. *Mensaje Bioquímico*, 42, 64-80. <http://bq.facmed.unam.mx/tab/wp-content/uploads/2020/02/7-Dra.-Sara-Frias.pdf>
- B. Reyes, P., García-de Teresa, B., Juárez, U., Pérez-Villatoro, F., Fiesco-Roa, M. O., Rodríguez, A., ... Frías, S. (2022). Fanconi anemia patients from an indigenous community in Mexico carry a new founder pathogenic variant in FANCG. *International Journal of Molecular Sciences*, 23(4), 2334. [doi:10.3390/ijms23042334](https://doi.org/10.3390/ijms23042334).
- C. Torres, L., Juárez, U., Reyes, P., & Frías, S. (2023). Anemia de Fanconi, Parte 2. Estrategia metodológica para el diagnóstico molecular en pacientes con anemia de Fanconi. *Acta pediátrica de México*, 44(1), 29–55. [doi:10.18233/apm44no1pp29-552548](https://doi.org/10.18233/apm44no1pp29-552548).



Memoria del 45º Taller de Actualización Bioquímica, Facultad de Medicina; UNAM

Origen y consecuencias de la inestabilidad genómica Síndromes de inestabilidad cromosómica

Origin and consequences of genomic instability Chromosomal instability syndromes

Juárez-Figueroa, Ulises^{1,2}; Ayala-Zambrano, Cecilia¹; Reyes, Pedro^{1,2} y Frías, Sara^{1*}

1. Laboratorio de Citogenética, Instituto Nacional de Pediatría/ Unidad de Genética de la Nutrición, Instituto de Investigaciones Biomédicas, UNAM.

2. Programa de Doctorado en Ciencias Biomédicas, UNAM.

*Correspondencia: Instituto Nacional de Pediatría, Torre de Investigación, 6º. Piso, Insurgentes Sur 3700-C, Col. Insurgentes Cuicuilco, Coyoacán, Ciudad de México. CP 04530 Tel +52(55) 1084-5533, sarafrias@biomedicas.unam.mx

Resumen

El DNA es nuestro patrimonio genético, es la única molécula que no puede reemplazarse a lo largo de toda la vida celular y se encuentra expuesto a diversos agentes que lo dañan, tanto endógenos como exógenos; se estima que una célula humana presenta entre 10^4 - 10^5 lesiones en el genoma por día, que pueden derivar en mutaciones, si no son restauradas, por lo que para mantener la homeostasis, nuestras células tienen toda una batería de mecanismos que se encargan de identificar, señalar y reparar las lesiones del DNA; se calcula que los genes que intervienen para mantener la integridad genómica son cientos y si fallan pueden ser letales o conducir a síndromes de inestabilidad genómica. Existen diversos tipos de lesiones y de mecanismos de reparación; las rupturas de doble hebra del DNA (DSB) son de las más deletéreas para la célula y de no repararse, o repararse erróneamente, se genera daño estructural y numérico en los cromosomas, resultando en una variedad de inestabilidad genómica que es la inestabilidad cromosómica y ésta genera un fenotipo celular y clínico anormal. Los síndromes de inestabilidad cromosómica más estudiados son la Ataxia telangiectasia, el Síndrome de Nijmegen, el Síndrome de Bloom y la Anemia de Fanconi, todos ellas enfermedades raras que tienen en común malformaciones congénitas, alteraciones inmunológicas, envejecimiento prematuro y un riesgo

Abstract

The DNA is our genetic heritage, it is the only molecule that can't be replaced throughout the cellular life and is exposed to both endogenous and exogenous agents that damage it; it is estimated that a human cell receives 10^4 - 10^5 lesions per day, which if they are not repaired, can lead to mutations, so to maintain cellular homeostasis, our cells have a battery of mechanisms that are responsible for identifying, signaling and repair DNA lesions; It is calculated that we have hundreds of genes involved in maintaining the genomic integrity; failure in one of them, may be lethal or lead to genomic instability syndromes. There are several types of DNA damage and repair mechanisms; DNA double-strand breaks (DSB) are one of the most deleterious lesions and if it is not repaired or if repaired erroneously, structural and numerical damage is generated in the chromosomes, resulting in a variety of genomic alterations leading to chromosomal instability generating an abnormal cellular and clinical phenotype. The most studied chromosomal instability syndromes are Ataxia telangiectasia, Nijmegen Syndrome, Bloom Syndrome and Fanconi Anemia, all of them are rare diseases that have in common congenital malformations, immunological alterations, premature aging and a very high risk of developing cancer. In addition, every of these syndromes shows phenotypic characteristics related to the specific gene defect. The study of these diseases has generated

muy elevado de desarrollar cáncer. Además de esto, cada uno de ellos tiene características particulares que se relacionan con el defecto génico específico. El estudio de estas enfermedades ha generado conocimiento sobre los mecanismos de reparación del DNA en humano, lo cual ha repercutido en un mejor diagnóstico, pronóstico y tratamiento de estos pacientes, pero también en el avance y en el conocimiento de la biología del cáncer y su tratamiento.

Palabras clave: inestabilidad genómica, reparación de DNA

knowledge regards the mechanisms of DNA repair in humans, which has resulted in a better diagnosis, prognosis and treatment of these patients, but also in the advancement in the knowledge of cancer biology and its treatment.

Key words: genomic instability, DNA repair

Introducción

Nuestro organismo posee células altamente complejas, tanto en estructura como en función, cada una de ellas contiene en su núcleo DNA asociado a proteínas histonas y no histonas que forman la cromatina, que se visualizará en la mitosis como 46 cromosomas, debido a la alta condensación generada (Figura 1). Estos cromosomas portan en conjunto un DNA con 3200 millones de pares de bases A-T y C-G, que constituyen nuestro genoma diploide y están ordenadas en una secuencia específica, y es donde se encuentra almacenada la información genética y las instrucciones necesarias para construir y mantener funcionando cada célula y el conjunto organizado de células que constituye un individuo; así mismo el DNA es el responsable de la transmisión de esta información de célula a célula mediante mitosis y de generación a generación mediante meiosis, que forma gametos haploides. Lo extraordinario de nuestro genoma, es que aunque tiene la capacidad para hacer que una célula pueda recambiar prácticamente todos sus tipos de componentes, los cromosomas no pueden recambiarse; esto tiene implicaciones muy importantes para la vida, puesto que si se daña el DNA necesita ser reparado, sin embargo, se ha observado que aún en un individuo normal, la reparación puede no ser perfecta, por lo que tiende a acumular lesiones y mutaciones que eventualmente conducirán a la pérdida celular, al envejecimiento o bien si hay mutaciones que confieran a una célula una ventaja reproductiva, puede conducir al desarrollo de cáncer [1].

Daño y reparación del DNA en homeostasis celular

Todo ser humano inicia su existencia como un cigoto, con 23 cromosomas de origen paterno y 23 cromosomas de origen materno para integrar un genoma propio de 46 cromosomas; este genoma se conserva extremadamente similar en todas las células

de un organismo adulto, a pesar de haber pasado por miles de millones de replications y de estar expuesto a múltiples agresiones endógenas y exógenas. Se estima que una célula humana presenta entre 10^4 - 10^5 lesiones en el genoma por día, inducidas por fuentes endógenas y exógenas y que pueden derivar en mutaciones, si no son restauradas [2-4].



Figura 1. Cromosomas humanos en metafase. Cada célula tiene 46 moléculas de DNA, visibles al microscopio como 46 cromosomas, cada uno contiene una molécula de DNA asociada a proteínas histonas y no histonas. Se aprecian los 46 cromosomas íntegros, sin alteraciones.

Fuentes endógenas de daño al DNA

Existen diferentes fuentes endógenas de daño al DNA entre las que destacan: a) la replicación del DNA, b) el contacto con productos del propio metabolismo celular y c) reacciones espontáneas intrínsecas a la naturaleza química del DNA [3].

a) Las células entran en ciclo celular para reproducirse, durante la replicación del DNA, las DNA polimerasas con frecuencia agregan bases nitrogenadas no complementarias; siendo el principal factor de mutaciones espontáneas en un organismo. Por lo general, la síntesis del DNA se lleva a cabo por las polimerasas de alta precisión α , δ y ϵ , que se estima generan en promedio un error por cada 10,000 bases incorporadas. Sin embargo, en ocasiones se requiere de la acción de polimerasas de síntesis translesión para completar la replicación. Estas polimerasas son menos precisas e introducen hasta 2 errores por cada 1000 bases ya que cuentan con un sitio catalítico de mayor tamaño que les permite acomodar bases modificadas o voluminosas que incluyen bases con alteraciones [1, 2, 4, 5].

b) Por otra parte, el DNA se encuentra situado en un ambiente acuoso en el que, con frecuencia, es

posible encontrar especies reactivas de oxígeno y nitrógeno, las cuales pueden dañar el DNA generando enlaces covalentes cruzados, rupturas de hebra sencilla, rupturas de doble hebra, modificación de bases, etc. (Cuadro 1, Figura 2), [6, 7]. La cantidad requerida de estas especies para causar daño varía, por ejemplo, en el caso del acetaldehído (Figura 2) [6].

c) Por último, todas las macromoléculas biológicas presentan descomposición espontánea, y el DNA no es la excepción. Las moléculas del DNA se unen entre sí por enlaces químicos, el enlace N-glicosídico formado entre la pentosa y la base nitrogenada es susceptible a hidrolizarse en condiciones fisiológicas, lo que implica la desaminación y pérdida de bases nitrogenadas causando cambios en las bases originales; por ejemplo la Citosina puede generar Uracilo, lo cual genera una mutación tipo transición de C>T [8, 9].

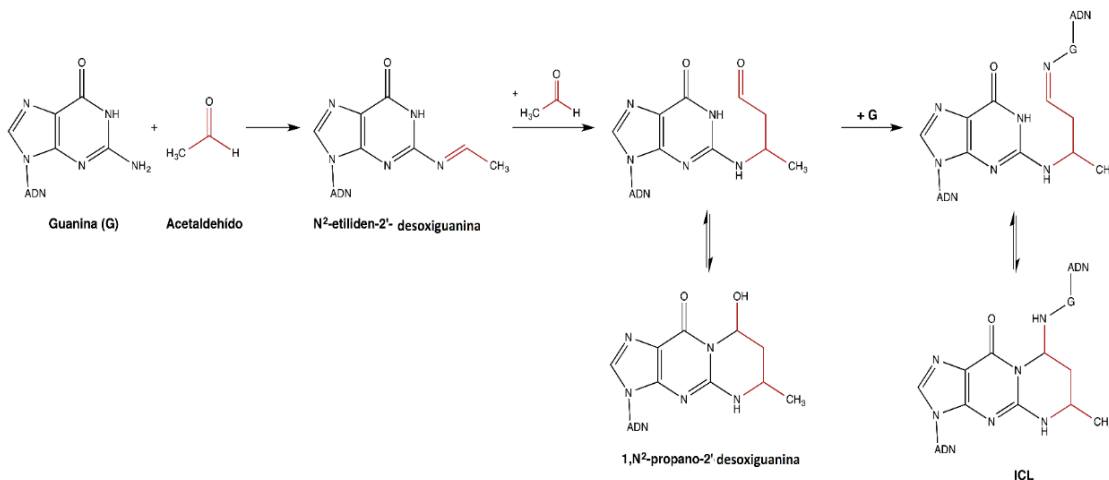


Figura 2. Mecanismo de generación de enlaces covalentes cruzados en el DNA, por acetaldehído.

Cuadro 1. Efecto de las fuentes de daño endógeno sobre el DNA.

Fuente de daño endógena	Ejemplo	Efecto en DNA
Especies reactivas de oxígeno (ROS, por sus siglas en inglés)	Peróxido de hidrógeno. Ion superóxido. Radical hidroxilo.	Oxidación de bases (8-oxo-deoxiguanosina). Ruptura de hebra sencilla. Ruptura de doble hebra.
Especies reactivas de nitrógeno	Óxido nítrico Peroxinitrito	Oxidación de bases (8-nitro-deoxiguanosina) Ruptura DNA Desaminación de guanina a adenina
Productos de la peroxidación de lípidos	Crotonaldehído Acroleína 4-hidroxinonenal (HNE) Malonaldehído (MDA)	Aduetos exocíclicos
Agentes alquilantes	Acetaldehído	Enlaces covalentes cruzados o Aduetos. Sustitución de bases ($\text{G} \rightarrow \text{A}$, $\text{A} \rightarrow \text{T}$).
Metabolismo del estrógeno	4-hidroxiestradiol	Aduetos. Modificación covalente de DNA.

Fuentes exógenas de daño al DNA

Una de las más importantes fuentes de daño al DNA en los seres vivos, es la radiación solar; existen tres categorías de radiación UV: A, B y C. Debido a que la capa de ozono absorbe UV-C, las principales fuentes de daño al DNA son la luz UV-A y UV-B, quienes pueden causar dos tipos de lesiones en el DNA: 1) la unión covalente de dos pirimidinas adyacentes, timina-timina o timina-citosina, formando dímeros ciclobutano pirimidina; 2) la formación de dímeros pirimidina-pirimidona o 6-4 fotoproductos entre bases de timina y citosina adyacentes [2, 6]. Otro agente físico que genera alteración en la estructura del DNA es la radiación ionizante (IR, por sus siglas en inglés), presente en el decaimiento de elementos radioactivos (como el uranio), la radiación cósmica, tratamientos médicos que utilizan rayos X o radioterapia, etc. La IR produce lesiones químicamente parecidas a las provocadas por las especies reactivas de oxígeno (ROS, por sus siglas en inglés), e incluyen más de 850 modificaciones de pirimidina, 450 de purinas, rupturas de hebra sencilla y rupturas de doble hebra; el daño producido por la IR no se distribuye homogéneamente y puede generar dos o más lesiones dentro de una o dos vueltas de la doble hélice [6, 10]. Sumado a estos eventos, la exposición

a factores químicos en consecuencia a nuestra forma de vida; como el fumar, el usos de pesticidas o agentes quimioterapéuticos, por ejemplo la mostaza nitrogenada, la mitomicina C y el cisplatino reaccionan químicamente con el DNA [9] y se genera una molécula con alteraciones que incluyen modificaciones químicas en las bases nitrogenadas, ruptura de una o ambas hebras y enlaces covalentes entre bases de cadenas complementarias, llamados enlaces covalentes cruzados o ICL (*interstrand cross link*) (Figura 3), [10].

Respuesta celular al daño en el DNA

Las fuentes exógenas y endógenas de daño al DNA, pueden provocar en él una amplia variedad de lesiones y para poder contrarrestar estos insultos, la célula tiene un mecanismo denominado respuesta al daño en el DNA o DDR (*DNA damage response*), compuesto por una red compleja y orquestada de interacción de diferentes vías de señalización, las cuales tienen la habilidad de a) censar el daño generado, b) transducir esta señal a la célula para reparar la lesión y c) modificar la progresión del ciclo celular; sin embargo, si el daño es excesivo, la célula puede tomar el camino de la senescencia o en último caso activar la apoptosis (Figura 4), [1, 10-12].

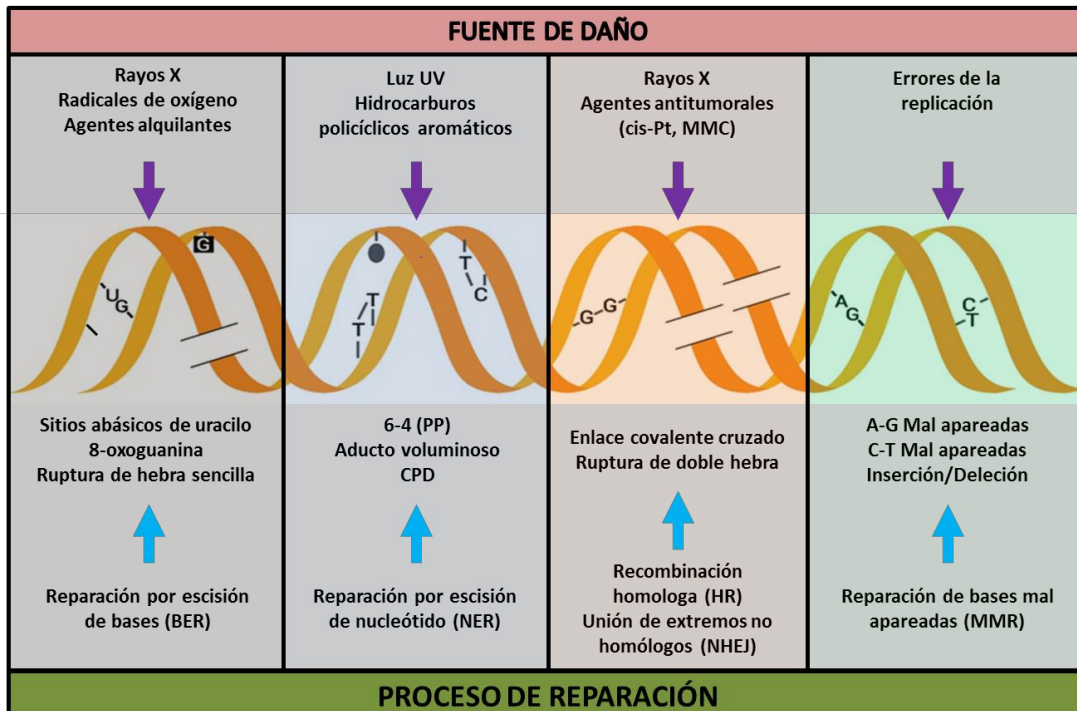


Figura 3. Fuentes de inducción de daño al DNA, lesiones generadas por ellas y mecanismos de reparación específicos.

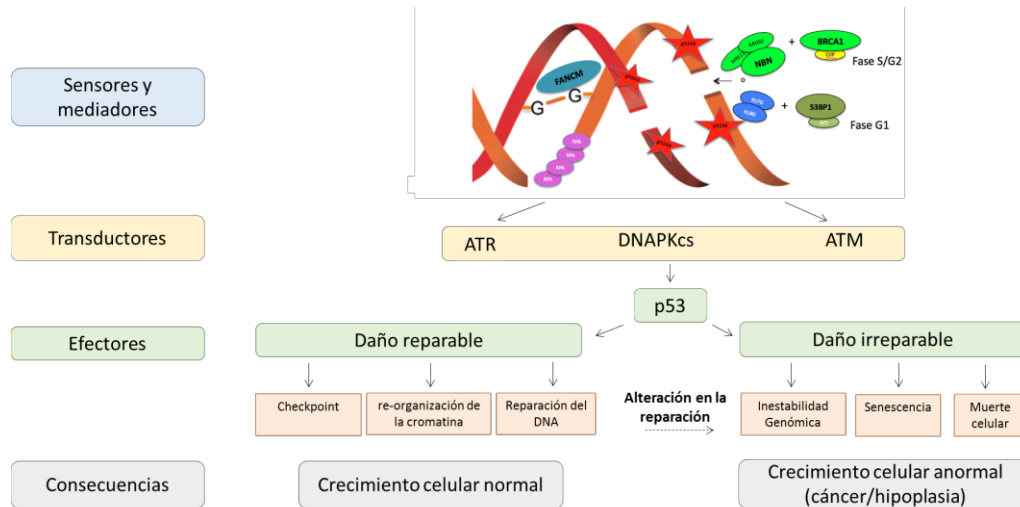


Figura 4. Mecanismos de acción de la DDR ante diversos tipos de lesiones en el DNA.

La DDR se realiza mediante el control de proteínas que participan como sensores, transductores y efectores de las lesiones generadas en el DNA, en este artículo revisaremos primordialmente aquellas que generan inestabilidad cromosómica.

a) La DDR se inicia una vez que los sensores reconocen la lesión, algunos de estos son el complejo MRN (Mre11–Rad50–Nbs1) que reconoce rupturas de doble hebra (DSBs, por sus siglas en inglés) en fase S/G2 y el heterodímero Ku70/Ku80 que actúa en fase G1; otro sensor muy importante es la proteína FANCM la cual reconoce horquillas de replicación estancadas por la presencia de un ICL [13, 14]

b) Una vez reconocida la lesión, se realiza una amplificación de la señal, las proteínas transductoras encargadas de este proceso son la familia PIKK cinasas (*Phosphatidy Inositol 3-Kinase-related Kinases*), compuesta por ATM, ATR y DNAPKcs [13, 14]. Estas cinasas actúan a través de modificaciones postraduccionales (por ejemplo, mediante fosforilación), de varias proteínas blanco. ATM y ATR fosforilan a CHK2 y CHK1 respectivamente, las cuales activarán un checkpoint del ciclo celular que lo detiene para dar paso a la reparación (Figura 5). Otras modificaciones se realizan sobre la cromatina, específicamente, estas PIKK cinasas pueden fosforilar a la variante de histona H2AX en el residuo Ser139 para dar γ H2AX, la cual marca el DNA con DSBs y genera el reclutamiento de otras proteínas para poder procesar la lesión y repararla [6, 15]. Existe otra serie de modificaciones postraduccionales de histonas canónicas, que facilitan el acceso a la cromatina para la detección y reparación del daño al DNA.

c) Por último, se activa un mecanismo específico de reparación dependiendo del daño generado; por ejemplo, en el caso de un ICL se pueden activar diferentes vías para realizar la reparación como la vía FA/BRCA, síntesis translesión (TS), reparación por escisión de nucleótidos (NER), recombinación homóloga (HR) o unión de extremos no homólogos (NHEJ); estas vías se discutirán a continuación [16, 17]. Una vez que la lesión se repara, las células continúan el ciclo celular por un proceso denominado “*checkpoint recovery*” [18, 19]. Sin embargo, si el daño no se repara, la célula puede 1. irse a senescencia, 2. realizar apoptosis o 3. dividirse con el daño no reparado, provocando inestabilidad cromosómica que puede derivar en disfunción celular y cáncer.

Por otro lado, en los últimos años se han dado grandes avances en el conocimiento de los procesos mediados por RNAs no codificantes (ncRNAs), y se han identificado microRNAs (miRNAs) que pueden regular la expresión de los RNAs mensajeros de las proteínas que participan en la DDR [20]. Otros ncRNAs que participan en este proceso, son los pequeños RNAs generados en la DDR (DDRNA) y los RNAs largos no codificantes inducidos por daño (dilncRNAs), y se ha observado que ambos tipos de RNA reguladores se producen después de que se genera una lesión, además de ser sitio-específicos y ayudar en el reclutamiento de factores de reconocimiento como el complejo MRN o 53BP1 [21].

Arresto del ciclo celular

La progresión del ciclo celular durante sus diferentes fases se controla principalmente por

complejos específicos ciclinas-cinasas dependientes de ciclina (CDK). Sin embargo, cuando una célula en división presenta daño en el DNA, se activa la DDR y el checkpoint propio de la fase en la que se encuentre la célula. Estos checkpoints son puntos donde la célula verifica que las actividades de cada fase se completen para poder proseguir con la siguiente. En caso de que exista daño en el DNA, la respuesta inmediata es la inhibición de los complejos ciclina-CDKs; como resultado, se genera un arresto del ciclo celular donde la célula puede realizar: a) reparación del daño o b) senescencia o apoptosis para prevenir la inestabilidad cromosómica, si el daño no puede ser reparado (Figura 5).

El mecanismo específico de respuesta para que el arresto celular se realice dependerá del tipo de lesión y de la fase del ciclo celular en la que se genere el daño, para así activar el “checkpoint” adecuado. Generalmente el no completar la replicación o debido a las lesiones del DNA provoca la activación de las PIKK cinasas ATM/ATR, que activan las cinasas CHK2/CHK1, respectivamente. CHK2/1, activa a la proteína p53, lo que resulta en la inducción de una amplia variedad de blancos transcripcionales, dentro de los cuales se encuentra la proteína p21 inhibidora

de CDK que se une e inhibe al complejo ciclina-CDK para bloquear la progresión del ciclo celular. Otro de los blancos activados por ATM es p38, una MAP cinasa que contribuye en la estabilización del mRNA que codifica a p21. Todo este proceso impide la progresión de G1 por la degradación de ciclina D (Figura 5a) o bien estos mismos actores pueden detener la transición de la fase G1 a S (Figura 5b); la progresión de fase S por daño al DNA o por alteración en la replicación (Figura 5c y 5d) y la transición de fase G2 a M (Figura 5e), en especial en la fase G2, además de la inhibición de CDC25A y CDC25B, la fosforilación de CDKs dependiente de WEE1, niveles bajos de p21 y la presencia de ATR y CHK1, son cruciales para el control y mantenimiento del “checkpoint” de fase G2, [22, 23]. Como se mencionó en el segmento anterior la reparación dependerá de la resección que tenga el DNA. En todos estos puntos interviene también la fosfatasa CDC25A, encargada de la reversión de la fosforilación inhibitoria de las CDK [22, 23]. Estos “checkpoints” que arrestan el ciclo celular, permiten a la célula tener tiempo para que reparen las DSBs, generalmente por NHEJ y por HR de manera post-replicativa durante las fases S y G2.

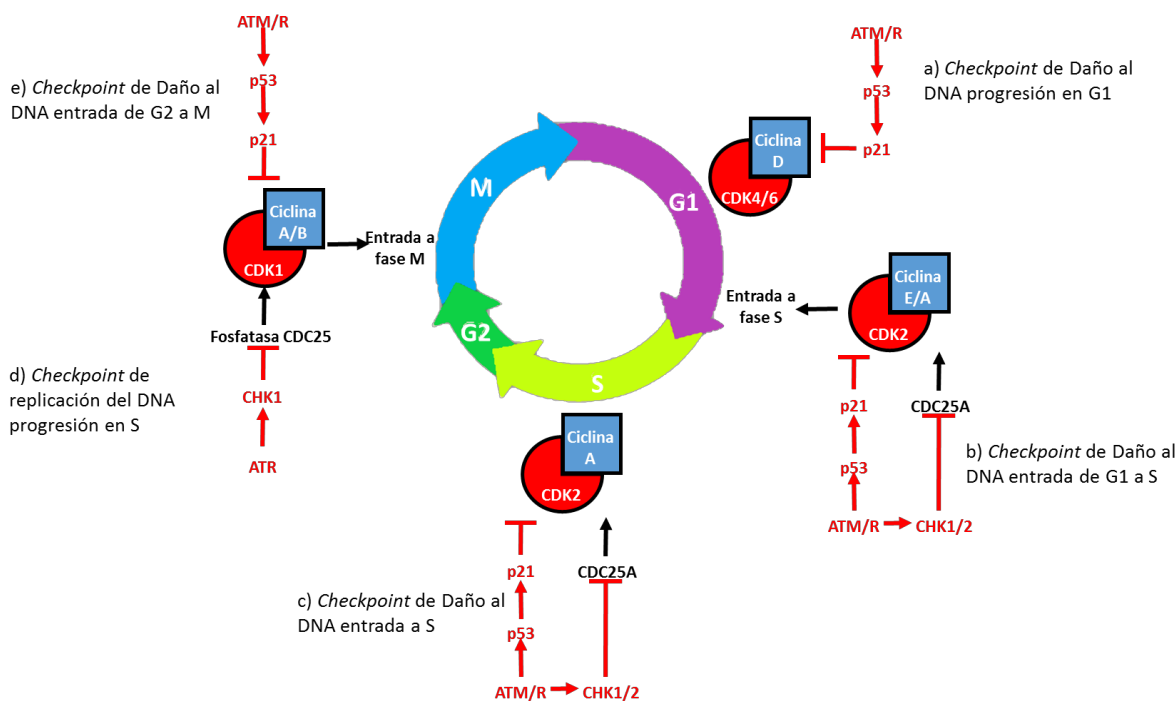


Figura 5. Inhibición de los complejos ciclinas-CDKs en la progresión del ciclo celular en respuesta a daño al DNA o alteración en su replicación.

Mecanismos de reparación del DNA

La reparación del DNA se considera como una de las herramientas que utiliza la célula como parte de la

DDR. Dos factores principales determinan la elección de la vía de reparación que será utilizada: el tipo de lesión presente en el DNA y la fase del ciclo celular en la cual son detectadas las alteraciones. El sistema de reparación de apareamientos erróneos (MMR, *mismatch repair*) se encarga de corregir bases mal apareadas; la reparación por escisión de bases (BER, *base excision repair*) se utiliza para corregir bases alteradas, mientras que la reparación por escisión de nucleótidos (NER, *nucleotide excision repair*) actúa sobre una variedad de lesiones que distorsionan la estructura helicoidal del DNA [1, 24]. Algunas lesiones como los ICLs, interfieren con procesos celulares indispensables, como la replicación y la transcripción, en este caso la vía que se encarga de su reparación es la FA/BRCA, que se considera como una vía maestra que procesa los ICLs y para repararlos, forma intermediarios tipo aductos que distorsionan el DNA y se reparan por NER y también forma DSBs. Las DSBs se consideran como una de las lesiones más peligrosas para la integridad del genoma y pueden originarse por una lesión directa o bien como producto de la reparación de un ICL [24]. Las DSBs se puede reparar por dos mecanismos y que dependen de la fase del ciclo en el que se encuentre la célula: 1) la recombinación homóloga, (HR) que opera en fase S/G2 y es dependiente de la presencia de una molécula homóloga de DNA, y 2) la unión de extremos no homólogos (NHEJ), que opera principalmente en G1 pero puede operar en cualquier fase del ciclo celular ya que no requiere de homología [24]. Una importante diferencia entre ambas vías, es que la HR se considera como un proceso fidedigno de reparación, mientras que la NHEJ es un proceso con tendencia a generar errores, sobre todo si se trata de la vía llamada NHEJ alterna (A-NHEJ, *alternative-non homologous end joining*), ya que puede ocasionar alteraciones cromosómicas o subcromosómicas que producen deleciones o duplicaciones [25, 26].

Las lesiones que pueden generar inestabilidad cromosómica, tema central del presente trabajo, son esencialmente las ICL y DSB; los procesos de reparación que intervienen para eliminarlas del DNA se presentan a continuación.

Reparación por escisión de nucleótidos o NER

Es la encargada de reparar lesiones que ocasionan la distorsión de la doble hélice como dímeros de ciclobutano-pirimidinas y los fotoproductos 6-4 pirimidina-piromidona, inducidos por luz UV; aquellas generadas por ciertos productos químicos como el cisplatino, así como por ciclopurinas derivadas de ROS [4]. NER presenta dos sub-rutas, la que se encarga de la reparación en el genoma global

(GG-NER, *global genome nucleotide excision repair*) y la que está acoplada a la transcripción (TC-NER, *transcription-coupled nucleotide excision repair*), que se activa ante lesiones en una hebra de DNA que servirá de template para la generación de RNA mensajero (Figura 6), [25].

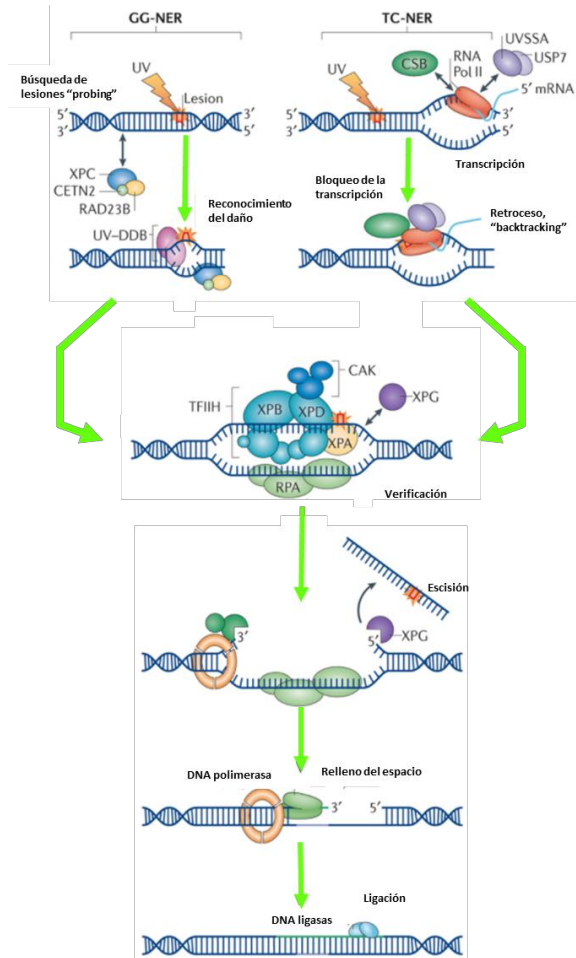


Figura 6. Mecanismo de reparación por escisión de nucleótidos (NER). GG-NER reparación global del genoma; TC-NER reparación acoplada a la transcripción.

El primer paso para ambas rutas es el reconocimiento del daño. En GG-NER, las proteínas relacionadas con el síndrome de Xeroderma pigmentosum son las principales involucradas; el complejo de las proteínas XPC/RAD23B/CETN-2 con ayuda de UV-DDB, reconoce la lesión que distorsiona el DNA; XPC permanece sobre la hebra opuesta a la lesión y su terminal carboxilo se inserta entre las hebras de DNA facilitando así la apertura de las hebras y posteriormente, RAD23B se disocia de la unión [4, 27]. En la sub-ruta TC-NER, el daño se reconoce durante la elongación del transcrito debido al estancamiento de la RNA Pol II en la lesión, en esta etapa la enzima interactúa con UVSSA, CSB y USP7

[4]. Una vez que la polimerasa localiza la lesión, la afinidad de CSB por la RNA Pol II aumenta y se recluta CSA, formándose el complejo CSA-CSB, que induce el retroceso (*backtracking*) de la RNA Pol II, dejando accesible el sitio de la lesión para ser reparado [27].

Los pasos posteriores son comunes para ambas sub-rutas. Así, el complejo TFIIH se recluta al sitio de la lesión. En seguida, se disocia el sub-complejo con actividad cinasa de TFIIH (CAK) y la doble hélice se separa gracias a la actividad helicasa del complejo, siendo la subunidad XPD la encargada de verificar la presencia de una lesión, con la colaboración de XPB y XPA [25]. Posteriormente, RPA se une a la hebra no dañada y XPA recluta al heterodímero XPF-ERCC1, el cual se dirige a la hebra lesionada y genera una incisión en su extremo 5'. Luego, XPG corta la hebra a unos 22-30 nucleótidos en dirección 3' de la lesión, que provoca la escisión del fragmento donde se encontraba la lesión. Para rellenar el espacio se recluta, mediado por PCNA, una DNA polimerasa delta o épsilon. El proceso culmina cuando una DNA ligasa I o III, genera el enlace covalente que unirá la hebra recién sintetizada con la hebra original que contenía la lesión (Figura 6), [4].

Reparación por recombinación homóloga (HR)

Como se mencionó anteriormente, la reparación de un DSB puede ser reparado por HR o NHEJ dependiendo de la fase en la que se encuentre la célula. Primero los extremos de DNA rotos son reseccionados por exonucleasas, con una mayor longitud de resección en la HR, y casi nula en la NHEJ [28]. En la RH se copia la información sin error que contiene la cromátida hermana no dañada y consta de cuatro pasos (Figura 7):

Señalización de la ruptura: protagonizada por las proteínas γ H2AX, MDC1 y RNF8 tras la activación de las cinasas ATM y ATR y la consecuente fosforilación de sus proteínas blanco CHEK1, CHEK2, TP53, BRCA1 y H2AX [29].

Resección de los extremos 5' de la DSB: el complejo MRN formado por MRE11, NBS1 y

RAD50, colabora con las exonucleasas CtIP, que genera una resección corta que se extiende por las exonucleasas EXO1 o DNA2, acopladas con las helicasas WRN o BLM [28, 30].

Invasión de la hebra: el DNA de cadena sencilla se recubre por RPA, la cual se intercambia por RAD51/FANCR con ayuda de BRCA2/FANCD2 y PALB2/FANCN para generar el filamento presináptico; éste se alinea por complementariedad con la hebra homóloga [17, 31].

Formación y resolución de intermediarios de Holliday: La síntesis del DNA faltante a partir de los extremos 3' provoca la formación de una estructura de cuatro brazos bicatenarios unidos entre sí (intermediarios de Holliday), los cuales se resuelven por helicasas y endonucleasas que, dependiendo donde lleva a cabo la incisión generan o no recombinación [32].

Debido a que la información se copia de la cromátida hermana, la reparación de las DSB por la RH es una vía fidedigna que se considera libre de error.

Reparación por unión de extremos no homólogos (NHEJ)

La reparación de DSB por la NHEJ, contrario al resultado de la reparación por HR, es propensa a error y no requiere de la presencia de una cromátida hermana para llevarse a cabo, por lo que la vía puede emplearse en cualquier fase del ciclo celular [28].

El inicio de la NHEJ requiere de la unión del heterodímero Ku70/80 en los extremos rotos del DNA, éste protege al DNA evitando su degradación y la actividad de las proteínas de la HR, además de servir como plataforma para reclutar proteínas de reparación como la DNA-PKcs. La DNA-PKcs es una cinasa con un cofactor proteico con actividad exonucleasa (Artemisa), el cual elimina algunos nucleótidos de los extremos rotos para que se pueda llevar a cabo la ligación por los complejos XRCC4/XLF y la DNA ligasa IV. Ésta resección, aunque pequeña, representa la modificación de la información en el sitio de daño, explicando los errores asociados a la reparación por NHEJ (Figura 7), [28, 33].

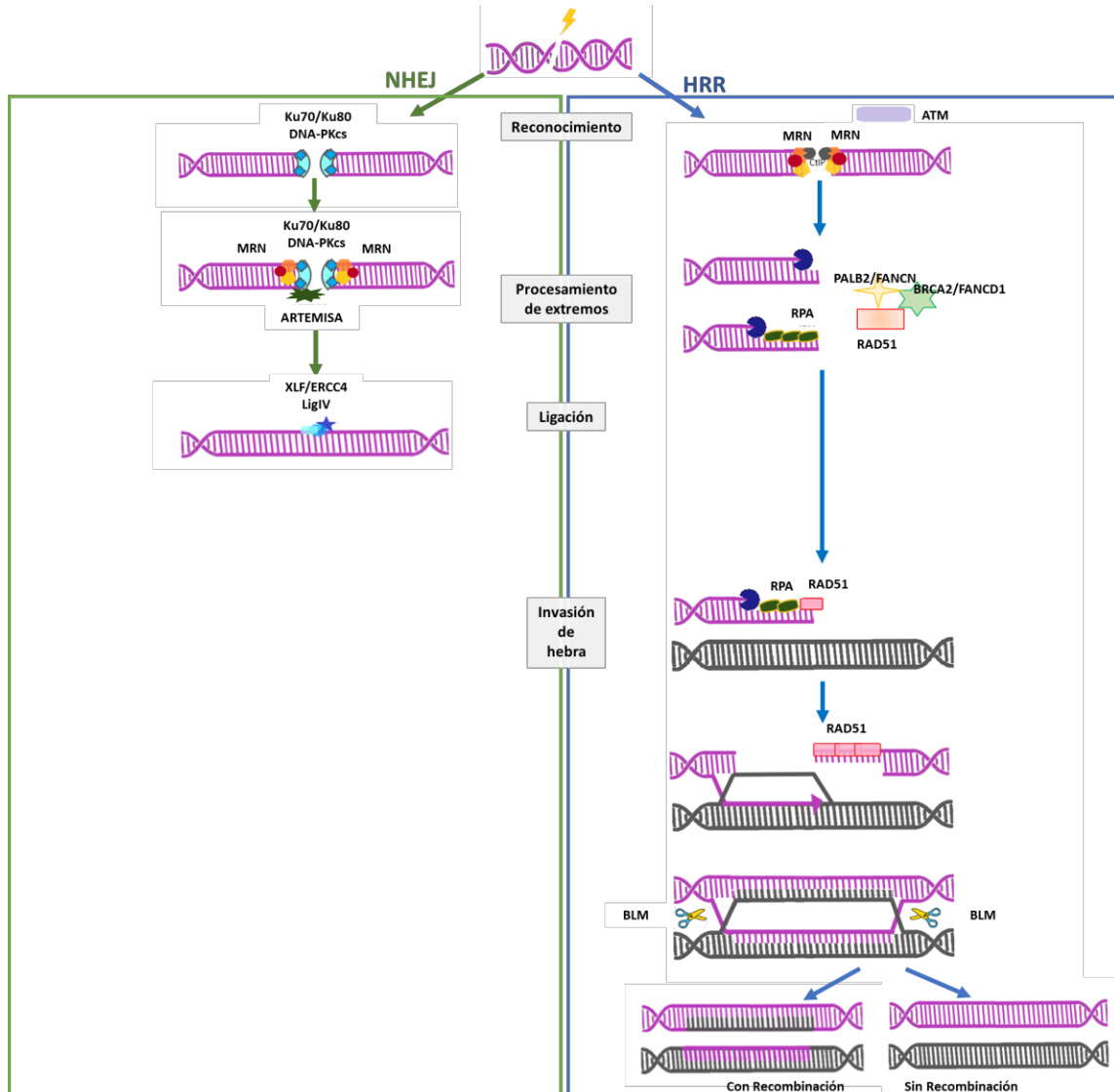


Figura 7. Reparación de DSB por NHEJ y HR. En la ligación durante la NHEJ, se pueden unir los extremos contiguos a la ruptura (como se muestra en la imagen) o extremos productos de rupturas de cromosomas distintos, lo que genera aberraciones cromosómicas como translocaciones, dicéntricos y figuras radiales.

Vía FA/BRCA

Los enlaces covalentes cruzados o ICLs, representan el daño más deletéreo que puede tener una célula, ya que esto provoca la inhibición de procesos esenciales como la transcripción y la replicación. Algunos agentes alquilantes procedentes de fuentes endógenas como el acetaldehído y formaldehído o exógenas como algunos fármacos como el cisplatino y la mitomicina C, reaccionan con las bases nitrogenadas formando un enlace covalente entre las hebras o ICL, éste no genera un cambio evidente en la estructura del DNA, por lo que su detección se llevará a cabo principalmente durante la fase S del ciclo celular, cuando el paso de la maquinaria de replicación

se vea interrumpido al no poder separar la doble hebra. El estancamiento de las dos horquillas de replicación genera la activación de la vía de reparación que se encarga de resolver este tipo de daño: la vía FA/BRCA [17, 34]. La vía FA/BRCA es compleja, ya que coordina diferentes mecanismos de reparación como NER, síntesis translesión y HR, así como la participación de diversas proteínas, entre las que destacan las 22 codificadas por los genes *FANCA*, *FANCB*, *FANCC*, *FANCD1*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *FANCL*, *FANCM*, *FANCN*, *FANCO*, *FANCP*, *FANCO*, *FANCR*, *FANCS*, *FANCT*, *FAUNCU*, *FANCV*, *FANCW* [35, 36].

Esta vía (Figura 8) inicia con el reconocimiento de la lesión por la proteína UHRF1 y el complejo FANCM-MHF1-MHF2, que reconocen estructuras ramificadas en el DNA generadas en horquillas estancadas y sirve como sitio de andamiaje para 11 proteínas, las proteínas río arriba de la vía FA/BRCA que incluye FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, FANCT, FAAP100, FAAP20 y FAAP24, en un complejo denominado “core AF” el cual tiene como función principal reclutar y ubiquitinar al heterodímero FANCD2-I; una vez posicionado y activado, FANCD2-I incorpora a las FANCP/SLX4, y las endonucleasas MUS81, SLX1 y XPF/ERCC4/FANCO quienes realizarán dos incisiones a ambos lados del ICL sobre una de las hebras parentales, liberando el ICL de una hebra y permitiendo el paso a la maquinaria de replicación, después de lo cual, se generan tres tipos de lesiones: una ruptura de hebra sencilla y un aducto en una de las

cromátidas y en la otra cromátida hermana se genera una DSB en donde las endonucleasas realizaron el corte. Cada una de estas lesiones se repara por diferentes vías de reparación (Figura 8), [34, 36]. La ruptura de hebra sencilla se repara por el complejo conformado por REV1, REV7/FANCV y REV3, el cual es una polimerasa de síntesis translesión capaz de incorporar nucleótidos tomando como molde la hebra complementaria con el aducto, este paso puede introducir mutaciones en la nueva hebra, debido a la baja fidelidad que tiene la polimerasa. El aducto se repara por vía NER y la DSB se procesa por las proteínas río abajo de la vía (FANCD1 / BRCA2, FANCN / PALB2, FANCS / BRCA1, FANCI / BRIP1, FANCO / RAD51C, FANCR / RAD51) que realizan la reparación por HR y así la lesión queda reparada sin errores para poder proseguir con el ciclo celular [34].

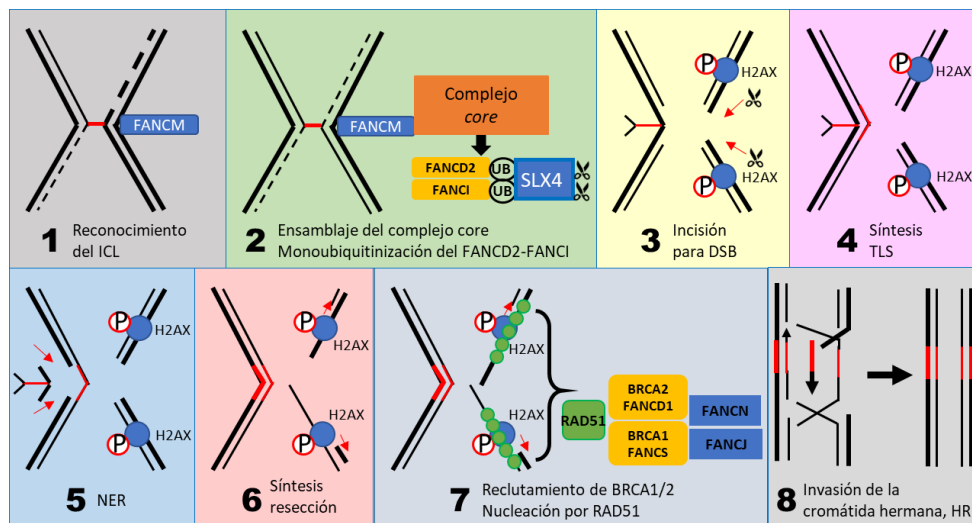


Figura 8. Vía FA/BRCA. En la figura observamos la secuencia de eventos del procesamiento de un ICL, desde el inicio donde se da el reconocimiento del ICL, pasando por el reclutamiento de las diferentes proteínas para el procesamiento de la lesión y por último el restablecimiento de las dos cromátidas hermanas integras.

Lesiones que causan alteraciones cromosómicas

Las alteraciones cromosómicas generalmente resultan de la ruptura y reunión de segmentos no homólogos, por lo que lo más común es que sean el resultado de la reparación no fidedigna de DSBs. El mal funcionamiento de alguno de los procesos de reparación (HR y NHEJ), genera mutaciones y aberraciones cromosómicas y cromatídicas, dependiendo de la fase del ciclo celular en la que se detecta la DSB [25]. Normalmente la HR no introduce errores, pero bajo algunas condiciones una sub-ruta de HR llamada reparación por alineamiento de cadena sencilla (SSA, *single strand annealing*) que repara

DSB flanqueados por secuencias de >25 nucleótidos repetidas (Figura 9), implica la pérdida de segmentos cromosómicos y puede generar translocaciones si el alineamiento se produce entre regiones con secuencias homólogas entre dos cromosomas diferentes [37]. Por otra parte, la conversión génica, resultado de la transferencia de información genética no recíproca entre dos secuencias de DNA, puede conducir a la formación de regiones con pérdida de heterocigidad (LOH, *loss of heterozygosity*) y también al intercambio físico entre dos segmentos cromosómicos (*crossing over*) que generan translocaciones, duplicaciones y deleciones cromosómicas (Figura 9), [26].

Sin embargo, en la mayoría de los casos de alteraciones cromosómicas estructurales, las proteínas que forman parte de la vía NHEJ canónica y alterna (C-NHEJ y A-NHEJ), son las responsables, puesto que ligan extremos que presentan homología en un segmento muy corto, de <25 nucleótidos, de manera que la NHEJ es capaz de unir de forma errónea e independientemente de su secuencia y estructura, cromosomas diferentes [37]. Por esta razón, si existe más de un DSB, se pueden inducir deleciones, inserciones o translocaciones, que pueden a su vez favorecer la formación de cromosomas dicéntricos u otros rearrreglos más complejos [38]. Se ha observado que en la variante alterna de la vía (A-NHEJ), independiente del dímero Ku 70/80 se incrementan los niveles de las uniones erróneas y translocaciones antes descritas (Figura 9), [39].

Alteración de la respuesta al daño al DNA. Síndromes de Inestabilidad cromosómica.

Generalidades

En un individuo sano, un porcentaje bajo de lesiones se convierten en mutaciones de manera a lo largo de nuestra vida y aún con sistemas de reparación normales, el DNA poco a poco se va deteriorando y conduce a envejecimiento.

Sin embargo, si el daño al DNA no se detecta o bien no se repara de manera normal, se genera un exceso de mutaciones tanto a pequeña escala, tipo mutaciones puntuales, como a gran escala, como mutaciones en el número o la estructura cromosómica, lo que genera inestabilidad genómica.

Generalmente, la inestabilidad genómica se produce cuando muta uno de los genes que intervienen en el mantenimiento de la integridad del DNA y esto origina diversas patologías en el humano, dependiendo del momento y el sitio de la mutación. Si ocurre en células somáticas dentro de un individuo normal, puede derivar en cáncer. Si la mutación ocurre en células germinales y se concibe un individuo con mutación constitucional, es decir en todas sus células, entonces se puede tener un individuo con un síndrome de inestabilidad genómica, con una alta y rápida acumulación de lesiones y mutaciones permanentes que pueden desembocar en muerte celular, la disminución de células dará falla en el crecimiento, deficiencia inmunológica o anemias. Pero si el daño celular es subletal, entonces se puede generar envejecimiento prematuro, y si eventualmente una célula adquiere ventaja reproductiva, se podría generar cáncer. Actualmente se reconocen más de 10 síndromes de inestabilidad genómica, entre ellos los llamados síndromes de inestabilidad cromosómica (SIC), porque en sus células se encuentran mutaciones a gran escala, visibles en los cromosomas (Cuadro 2).

Cuadro 2. Síndromes de inestabilidad cromosómica.

Enfermedad	Número probable de genes	Genes identificados	Proceso molecular alterado
Ataxia Telangiectasia	1	ATM	Respuesta y señalización de daño al DNA Tipo DSB
Síndrome Nijmegen	1	NBN	Reparación de DSB (reparación NHEJ)
Síndrome de Bloom	1	BLM (RECQL3)	Reparación de DSB, Desenrollamiento del DNA (reparación/replicación/recombinación)
Anemia de Fanconi	22	FANCA-FANCW	Reparación de enlaces covalentes intercatenarios (HRR), DSBs

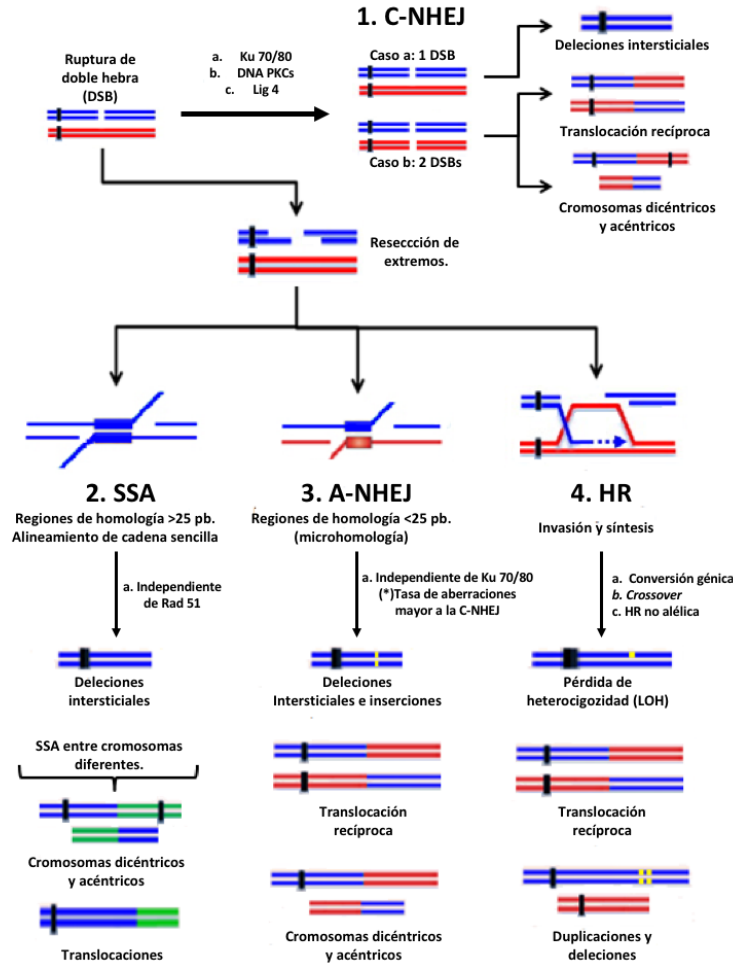


Figura 9. Mecanismos de reparación de DSB que conducen a rearrreglos cromosómicos.

Ataxia Telangiectasia (AT)

La AT tiene una incidencia de ~1/100,000 nacidos vivos. Presenta una herencia autosómica recesiva, se conoce un solo gen responsable, el *ATM* (*ataxia telangiectasia mutated*), localizado en 11q22.3-23.1 [40]. la mayoría de los pacientes presentan mutaciones puntuales patogénicas del tipo inserciones/deleciones intragénicas, mutaciones sin sentido y de sentido erróneo, aunque un pequeño porcentaje, <10% pueden presentar grandes deleciones y duplicaciones de exones completos o de todo el gen. Codifica para una proteína de alto peso molecular, la ATM que pertenece al grupo de las proteínas PIKs, cinasas que fosforilan residuos de serina/ treonina en un amplio espectro de sustratos relacionados con el reconocimiento de las DSBs y la respuesta celular a éstas; usualmente se encuentra como un dímero inactivo que se recluta por los DSBs mediante el complejo MRN, específicamente interactuando con el dominio carboxi-terminal de la proteína NBS1

(NBN). ATM se reconoce también como activadora de vías de señalización que coordinan los *check points*: G1-S fosforilando p53 en Ser15; G2/M fosforilando Chk2 en Thr68; y el *check point* intra-S fosforilando SMC1 en Ser957 y Ser966. Tiene también actividades citoplásmicas como respuesta al estrés oxidante, función mitocondrial, así como en la apoptosis (Figura 4), [40].

Las células de los pacientes con AT presentan inestabilidad cromosómica espontánea e hipersensibilidad a la radiación ionizante y a la bleomicina. Aunque pueden presentar rupturas cromatídicas y cromosómicas que son la evidencia directa de las lesiones tipo DSBs, lo característico de las células AT es que se encuentren translocaciones en sitios cromosómicos en los que se encuentran genes de inmunoglobulinas y receptores de ellas, específicamente translocaciones e inversiones en los cromosomas 7 y 14 en las bandas G 7p14, 7q35, 14q11-q12 y 14q32 (Figura 10), [41].

A nivel clínico, el fenotipo presenta una amplia variabilidad, pero se aprecia una correlación genotipo-fenotipo muy constante; cuando la mutación tiene como consecuencia la ausencia de la proteína ATM, el paciente presenta la forma clásica de la AT, con una presentación temprana y grave, mientras que si existe una mutación que permite la existencia de una actividad residual de ATM, se presenta una forma atenuada de la AT, con una presentación tardía y un fenotipo más leve.

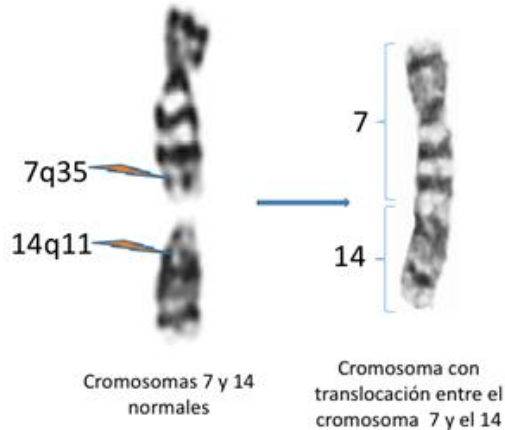


Figura 10. Translocación (7;14) (q35; q11). El rompimiento de los cromosomas, resultado de DSBs y su alteración en la reparación, trae como consecuencia la formación de translocaciones y otras alteraciones como deleciones o inversiones, son características de los Síndromes de Ataxia telangiectasia por mutación en el gen *ATM* y Síndrome de Nijmegen, por mutación en el gen *NBN*.

Las manifestaciones clásicas de la AT son telangiectasias oculocutáneas, degeneración neuronal progresiva, inmunodeficiencia con bajas concentraciones de IgA, IgE e IgG2 y alta susceptibilidad al cáncer.

Las manifestaciones neuronales incluyen ataxia cerebelosa progresiva; anomalías oculomotoras, particularmente apraxia ocular; trastornos del movimiento, tales como corea; y disfunción cognitiva. En la forma clásica, las manifestaciones pueden iniciar a los dos años y frecuentemente a los 10 años los pacientes pueden requerir silla de ruedas [40, 41]. La AT se presenta con afectación multisistémica, que incluye inmunodeficiencia, infecciones sinopulmonares, radiosensibilidad, telangiectasia oculocutánea, niveles elevados de alfa2-microglobulina sérica. Presentan envejecimiento prematuro con manifestaciones clínicas propias del adulto mayor como diabetes, hiperlipidemia, osteoporosis etc. y predisposición al cáncer. En la forma atenuada de AT, presenta actividad residual de ATM, ausencia de telangiectasias, funciones pulmonar y endocrina

normal, inmunoglobulinas normales y menor sensibilidad a la radiación ionizante; los pacientes tienen una esperanza de vida mayor, algunos pueden llegar a los 50 años y el cáncer aparece más tardíamente [42].

Los pacientes AT tienen 100 veces más riesgo que la población general de desarrollar algún cáncer y en edades más tempranas. Los más frecuentes son los linfomas, 50% tipo no Hodgkin, aunque también se presentan los Hodgkin en un 10%. Las leucemias linfoblástica aguda y linfocítica crónica representan el 25% y el resto son tumores sólidos como sarcomas, gonadoblastoma, meduloblastoma, cáncer de mama entre otros [41].

Es importante también hacer notar que los individuos portadores de la mutación, es decir heterocigotos con un alelo normal y un alelo *ATM* mutante, presentan una alta predisposición a cáncer de mama y otros tipos de cáncer, su riesgo es hasta 4 veces mayor que el de la población general [41].

Síndrome de Nijmegen (NBS)

El síndrome de Nijmegen o NBS (por sus siglas en inglés Nijmegen Breakage Syndrome), es una enfermedad rara, sin una prevalencia estimada hasta ahora, aunque se han reportado alrededor de 150 en la literatura médica, se sabe que hay muchos más pacientes en los registros internacionales [43]. Su forma de herencia es autosómica recesiva, originalmente se clasificó como una variante de AT, pero pronto se encontró que es otra entidad y se debe a la mutación de un gen originalmente llamado *NBS*, aunque posteriormente se cambió a *NBN*. La mutación hipomórfica más común es la deleción de 5 pares de bases en el exón 6, que produce un fragmento de proteína NBN de 26 kD y una proteína de 70 kD que se origina por activarse un codón de inicio río-arriba del codón normal. La nibrina o NBN es una proteína que normalmente forma el heterotrímero MRN compuesto por MRE11, RAD50 y NBN (Figura 4), el cual se requiere para fosforilar y así activar a las PIK cinasas ATR y ATM, por lo que su actividad es de primordial importancia en el reconocimiento de lesiones, la reparación del DNA y la activación de los *checkpoints* del ciclo celular; importantemente funciona como sensor para el reconocimiento y procesamiento de los DSBs, tanto en NHEJ como en HR (Figuras 4 y 7). En células murinas, los mutantes nulos para *nbn* tienen alterados los *checkpoints* de G2 e intra-S, muy similar a lo que ocurre con los mutantes *Atm*. La NBN tiene un dominio de unión con ATM, esta interacción activa ATM y le permite realizar la monomerización y autofosforilación necesarios para

el procesamiento de los DSBs. Esta cercana interacción explica las similitudes en el fenotipo clínico y celular entre AT y NBS [43]. A nivel celular, los pacientes con NBS presentan inmunodeficiencia celular y humoral, alteración en los *checkpoints* del ciclo celular e inestabilidad cromosómica espontánea e inducida por radiación ionizante y bleomicina, que se presenta como alteraciones rupturas cromosómicas y translocaciones que involucran los cromosomas 7 y 14, en 4 sitios frecuentes: 7p13, 7q35, 14q11 and 14q32, esto similar que en la AT (Figura 10).

Clínicamente, los pacientes con NBS presentan microcefalia severa que se manifiesta desde el nacimiento y progresa con la edad, tienen retraso leve del crecimiento, malformaciones de cerebro, de pulgares y genito-urinarias; inmunodeficiencia celular y humoral que predispone a infecciones recurrentes. Presentan manifestaciones de envejecimiento prematuro como encanecimiento temprano, insuficiencia ovárica prematura y un riesgo muy alto de desarrollar cáncer a una edad temprana; los principales tipos de cáncer son linfomas de todos los tipos y leucemias agudas linfoblásticas y mieloblásticas, aunque también presentan tumores sólidos como neuroblastomas y sarcomas [41, 43].

Síndrome de Bloom (SB)

El síndrome de Bloom es una enfermedad rara, para la cual no hay una incidencia establecida, tiene un patrón de herencia autosómica recesiva y se sabe que hay una alta incidencia en poblaciones endogámicas en donde la frecuencia de portadores se ha estimado tan alta como 1% [44]. El SB se debe a la mutación bialélica del gen *BLM*, localizado en 15q26.1. La mutación más común se encuentra en la población judía Ashkenazi, es una delección de seis nucleótidos en la posición 2281 y su reemplazo con otros siete nucleótidos. La proteína BLM es una helicasa de la familia RecQ, normalmente funciona restableciendo las horquillas de replicación dañadas y cooperando con la reparación de DSBs. Su función específica es realizar el desenrollamiento de la doble hélice, de manera que es una proteína multifuncional que interviene en el mantenimiento de los telómeros, la reparación por HR (Figura 7), tanto en las células somáticas, en donde coopera en la reparación de DSBs inducidos, así como programados para la formación de inmunoglobulinas. En las células germinales también es esencial para realizar la recombinación meiótica ya que cuando dos moléculas de DNA recombinan, se mantienen unidas por estructuras llamadas de Holliday, que se deben resolver para que pueda continuar el proceso de la recombinación. BLM tiene dos sitios de fosforilación, T99 and T122, la primera

relacionada con la supresión de intercambios de cromátidas hermanas (ICHs) y la segunda que impide que se formen alteraciones cromosómicas tipo figuras radiales [45], por lo que la falla de estas modificaciones tiene como consecuencia alteraciones cromosómicas.

A nivel celular, los pacientes con SB presentan un incremento de aberraciones cromosómicas espontáneas. Entre las más comunes son figuras radiales, que resultan por la falla en la reparación del DNA, dando lugar a la unión anormal de las cromátidas de cromosomas homólogos o no homólogos que forman estructuras de cuatro brazos llamados tetraradios (Figura 11), otra alteración cromosómica que caracteriza a SB, es un gran incremento de aproximadamente 10 veces más ICHs que las células normales (Figura 12), éstas características celulares se aprovechan para hacer el diagnóstico citogenético de SB [44, 45]



Figura 11. Figura de intercambio de cromátidas no hermanas. Se observa una figura radial tipo tetraradio (Flecha), característica de las células de pacientes con Síndrome de Bloom.

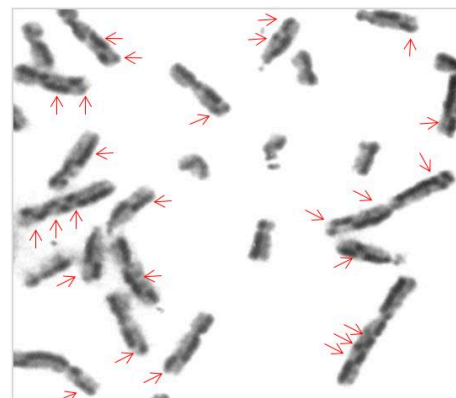


Figura 12. Imagen parcial de una célula de un paciente con Síndrome de Bloom. Se observan cromosomas con un exceso de intercambio de cromátidas hermanas o ICHs (flechas), resultado de la mutación en BLM. Una célula normal presenta entre 1-5 ICHs en total.

A nivel clínico, el SB la característica principal es una falla generalizada del crecimiento, que inicia desde *in útero*, continúa después del nacimiento y se manifiesta como un enanismo proporcionado, aunque el cráneo suele ser más pequeño y estrecho, con pabellones auriculares grandes; presentan inmunodeficiencia y una voz aguda. Tienen hipersensibilidad a la luz solar, por lo cual pueden presentar eritema telangiectásico, en piel también presentan áreas hipo e hiperpigmentadas (Figura 13), [1, 41, 44, 45]. Al igual que los otros síndromes de inestabilidad cromosómica, los pacientes con SB tienen un riesgo de desarrollar cáncer hasta 300 veces mayor que la población general, las neoplasias más comunes en ellos son la leucemia, linfoma y tumores sólidos como cáncer colorrectal, de laringe, mama y piel [44].



Figura 13. Paciente con síndrome de Bloom. Se observan las telangiectasias oculares (también presentes en Ataxia Telangiectasia) y el eritema por hipersensibilidad a la luz solar.

Anemia de Fanconi (AF)

La anemia de Fanconi es el síndrome de falla medular hereditario más frecuente, con una incidencia aproximada de 1-5 por millón de nacidos vivos, y una frecuencia de portadores de 1 en 300, aunque por tener un patrón de herencia primordialmente autosómico recesivo, en poblaciones endogámicas la frecuencia de portadores puede ser hasta de 1 en 77 [46]. La AF tiene alta heterogeneidad genética y la enfermedad se presenta cuando existe una mutación en cualquiera de los 22 genes *FANC* identificados hasta ahora: *FANCA*, *FANCB*, *FANCC*, *FANCD1*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *FANCL*, *FANCM*, *FANCN*, *FANCO*, *FANCP*, *FANCO*, *FANCR*, *FANCS*, *FANCT*, *FAUNCU*, *FANCV*, *FANCW* [35, 36], 20 de ellos tienen un patrón de herencia autosómico recesivo, *FANCB* recesiva ligada al X y *FANCR* autosómica dominante. Las proteínas que codifican se integran en una vía llamada FA/BRCA que opera primordialmente en fase S del

ciclo celular y se encarga de la identificación y reparación de los ICLs (Figura 8).

Las mutaciones en el gen *FANCA* son las más frecuentes, seguidas por las de *FANCC* y *FANCG* y juntos suman cerca del 90% de los pacientes y el resto se distribuyen entre los otros 19 genes [34]. Las mutaciones más frecuentes son las deleciones de *FANCA*, que representan el 30% y el resto son mutaciones prácticamente privadas, ya que hay muy pocas mutaciones fundadoras [1, 34].

Cuando alguno de los genes *FANC* está mutado, la reparación por HR que es libre de error no puede realizarse, por lo que en los pacientes AF se canalizan las lesiones en el DNA hacia la vía NHEJ, que es propensa a introducir errores en el genoma.

A nivel celular, se pueden identificar cuatro características principales: 1) Arresto en la fase G2 del ciclo celular, 2) Alta frecuencia de aberraciones cromosómicas espontáneas e inducidas por agentes alquilantes como el diepoxibutano y la mitomicina C, 3) hipersensibles a especies reactivas de oxígeno y 4) son proapoptóticas. El fenotipo celular es extremadamente constante, por lo que se utiliza su hipersensibilidad a los agentes alquilantes para realizar el diagnóstico de certeza de AF, mediante la prueba de fragilidad cromosómica inducida por mitomicina C o diepoxibutano, con lo cual se induce una frecuencia de aberraciones tipo rupturas cromatídicas, que representan las lesiones que llegaron a DSBs y ya no se repararon, y por otro lado hay una alta cantidad de figuras radiales, resultado de la reparación propensa a error que liga dos cromosomas diferentes (Figura 14), el incremento en aberraciones cromosómicas es de 10 o más veces con respecto a las células normales [41, 46-48].

A nivel clínico, el fenotipo es altamente variable, los pacientes con AF tienen manifestaciones que pueden identificarse en diferentes momentos de la vida:

- 1) Malformaciones congénitas, presentes en 75% de los pacientes [1, 41, 47, 48], son detectables desde el nacimiento (algunas prenatalmente), y comprenden por orden de frecuencia: alteraciones esqueléticas principalmente las que afectan el eje radial, alteraciones de pigmentación en la piel como hiperpigmentación y manchas café con leche, talla baja, hipoplasia o ausencia de radio, hipoplasia, ausencia o duplicación de pulgares, hipoplasia de eminencia tenar; malformaciones renales, microcefalia, cara triangular. Estas malformaciones pueden sobrelaparse con las que definen la asociación

VACTERL-H, acrónimo de (en inglés: *Vertebral anomalies, Anal atresia, Congenital heart disease, Tracheo-esophageal fistula, Esophageal atresia, Renal, Limb anomalies, and Hydrocephalus*) hasta en 33% de los pacientes AF [49].

- 2) Falla de la médula ósea, esta manifestación se presenta en más del 90% de los pacientes y es detectable en promedio a los 7.6 años, con disminución de todos los tipos celulares sanguíneos, es decir pancitopenia; muchos de los pacientes inician con trombocitopenia, seguida de neutropenia y después falla medular completa, por lo que presentan palidez, sangrados frecuentes e infecciones de repetición [46-48].
- 3) Predisposición a cáncer, los pacientes tienen un riesgo hasta de 500 veces mayor que la población general de desarrollar neoplasias. El cáncer se presenta tempranamente, en la adolescencia pueden iniciar comúnmente con leucemia mieloide aguda y posteriormente alrededor de los 30 años, se presentan tumores sólidos de todos tipos, pero el cáncer de cabeza y cuello es sin duda uno de los más frecuentes en AF [46, 47].

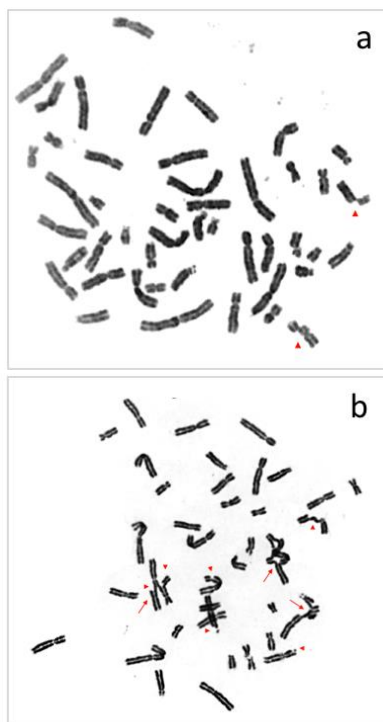


Figura 14. Aberraciones cromosómicas estructurales inducidas por un agente que produce enlaces covalentes cruzados. a) de un individuo normal, b) de un paciente AF, ambos tratados con diepoxibutano. Las puntas de flecha señalan rupturas cromatídicas, las flechas indican figuras radiales. Las figuras radiales son extremadamente raras en el individuo normal, mientras que son frecuentes en los pacientes con AF.

No existe una clara correlación entre el genotipo y el fenotipo de los pacientes AF, pero de manera general se observa que quienes portan mutaciones en genes río abajo de la vía FA/BRCA (es decir después de la función de FANCD2I ubiquitinizada), que abarca la reparación por HR, presentan una mayor incidencia de cáncer [1, 17, 47].

Finalmente, en AF también se presentan enfermedades propias de la tercera edad, tal como diabetes, menopausia precoz, dislipidemia, entre otros [1].

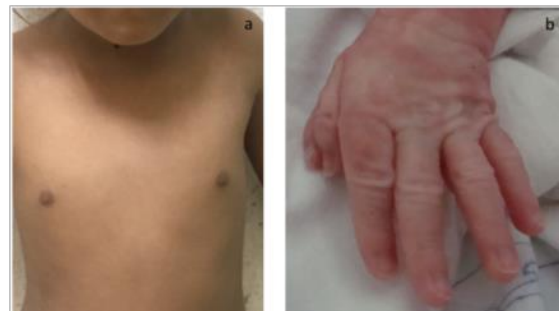


Figura 15. Manifestaciones clínicas en anemia de Fanconi. a) alteraciones en la pigmentación de la piel, b) Duplicación del pulgar en un recién nacido (Fotografía cortesía del Instituto Nacional de Perinatología).

Conclusiones

El contar con las apropiadas vías de reparación de diversas lesiones del DNA, mantiene el genoma íntegro a pesar de la exposición continua a agentes genotóxicos y a los millones de millones de divisiones celulares que se realizan a lo largo de la vida de un individuo.

La falla en los mecanismos de detección y reparación del daño al DNA, como hemos visto, genera un fenotipo especial que da lugar a estos síndromes y que tienen características en común como: falla en el crecimiento, malformaciones congénitas, la alta predisposición a cáncer y envejecimiento prematuro; todas ellas relacionadas con la acumulación de daño no reparado o mal reparado, que puede ser letal para la célula y generar hipoplasia o malformaciones congénitas y en caso de ser daño subletal, puede generar acumulación de lesiones y dar lugar a envejecimiento prematuro y cáncer (Figura 4).

Finalmente, para la generación de conocimiento y para la atención médica, el estudio de los individuos con síndromes de inestabilidad cromosómica es importante, ya que son modelos biológicos naturales

por los cuales se ha generado enorme cantidad de conocimiento sobre los mecanismos de reparación en humano, lo que ha permitido mejorar el diagnóstico, pronóstico y tratamiento de los individuos afectados y también de una diversidad de neoplasias en las que estos genes están mutados.

Referencias

- García-de, et al., *DNA Damage as a Driver for Growth Delay: Chromosome Instability Syndromes with Intrauterine Growth Retardation*. BioMed Research International, 2017. **2017**: p. 14.
- Clancy, S. *DNA damage & repair: mechanisms for maintaining DNA integrity*. Nature Education 2008.
- Hoeijmakers, J.H.J., *MOLECULAR ORIGINS OF CANCER DNA Damage, Aging, and Cancer*. New England Journal of Medicine, 2009. **361**(15): p. 1475-1485.
- Marteijn, J.A., et al., *Understanding nucleotide excision repair and its roles in cancer and ageing*. Nature Reviews Molecular Cell Biology, 2014. **15**(7): p. 465-481.
- Ohashi, E., et al., *Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the product of the human DINB1 gene*. Journal of Biological Chemistry, 2000. **275**(50): p. 39678-39684.
- Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease*. Nature, 2009. **461**(7267): p. 1071-1078.
- Voulgaridou, G.P., et al., *DNA damage induced by endogenous aldehydes: Current state of knowledge*. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 2011. **711**(1-2): p. 13-27.
- De Bont, R. and N. van Larebeke, *Endogenous DNA damage in humans: a review of quantitative data*. Mutagenesis, 2004. **19**(3): p. 169-185.
- Cheung-Ong, K., G. Giaever, and C. Nislow, *DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology*. Chemistry & Biology, 2013. **20**(5): p. 648-659.
- Ciccía, A. and S.J. Elledge, *The DNA Damage Response: Making It Safe to Play with Knives*. Molecular Cell, 2010. **40**(2): p. 179-204.
- Huang, Y.L. and L. Li, *DNA crosslinking damage and cancer - a tale of friend and foe*. Translational Cancer Research, 2013. **2**(3): p. 144-154.
- Hanawalt, P.C., *Historical perspective on the DNA damage response*. DNA Repair, 2015. **36**: p. 2-7.
- Cimprich, K.A. and D. Cortez, *ATR: an essential regulator of genome integrity*. Nature Reviews Molecular Cell Biology, 2008. **9**(8): p. 616-627.
- Sirbu, B.M. and D. Cortez, *DNA Damage Response: Three Levels of DNA Repair Regulation*. Cold Spring Harbor Perspectives in Biology, 2013. **5**(8).
- Rodríguez, A., et al., *Fanconi anemia cells with unrepaired DNA damage activate components of the checkpoint recovery process*. Theoretical Biology and Medical Modelling, 2015. **12**.
- Clauson, C., O.D. Scharer, and L. Niedernhofer, *Advances in Understanding the Complex Mechanisms of DNA Interstrand Cross-Link Repair*. Cold Spring Harbor Perspectives in Biology, 2013. **5**(10).
- Ceccaldi, R., P. Sarangi, and A.D. D'Andrea, *The Fanconi anaemia pathway: new players and new functions*. Nat Rev Mol Cell Biol, 2016. **advance online publication**.
- Medema, R.H. and L. Macurek, *Checkpoint recovery in cells: how a molecular understanding can help in the fight against cancer*. F1000 Biol Rep, 2011. **3**(10): p. 3.
- Halim, V.A., et al., *Comparative Phosphoproteomic Analysis of Checkpoint Recovery Identifies New Regulators of the DNA Damage Response*. Science Signaling, 2013. **6**(272).
- He, M.Y., et al., *MicroRNAs, DNA Damage Response, and Cancer Treatment*. International Journal of Molecular Sciences, 2016. **17**(12).
- Michellini, F., et al., *Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks*. Nature Cell Biology, 2017. **19**: p. 1400.
- Shaltiel, I.A., et al., *The same, only different - DNA damage checkpoints and their reversal throughout the cell cycle*. Journal of Cell Science, 2015. **128**(4): p. 607-620.
- Barnum, K.J. and M.J. O'Connell, *Cell Cycle Regulation by Checkpoints*. Cell Cycle Control: Mechanisms and Protocols, 2nd Edition, 2014. **1170**: p. 29-40.
- D'Andrea, A., *4 - \ DNA\ Repair Pathways and Human Cancer*. 2015: p. 47 - 66.e2.
- Thoms, K.-M.a.K.C.a.E.S., *Lessons learned from DNA repair defective syndromes*. Experimental Dermatology, 2007. **16**(6): p. 532--544.
- Kasperek, T.R. and T.C. Humphrey, *DNA double-strand break repair pathways, chromosomal rearrangements and cancer*. Seminars in Cell & Developmental Biology, 2011. **22**(8): p. 886-897.
- Tornaletti, S., D. Reines, and P.C. Hanawalt, *Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA*. Journal of Biological Chemistry, 1999. **274**(34): p. 24124-24130.
- Ranjha, L., S. Howard, and P. Cejka, *Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes*. Chromosoma, 2018.
- Walsh, C.S., *Two decades beyond BRCA1/2: Homologous recombination, hereditary cancer risk and a target for ovarian cancer therapy*. Gynecologic Oncology, 2015. **137**(2): p. 343-350.
- Liu, T. and J. Huang, *DNA End Resection: Facts and Mechanisms*. Genomics Proteomics & Bioinformatics, 2016. **14**(3): p. 126-130.
- Colavito, S., R. Prakash, and P. Sung, *Promotion and regulation of homologous recombination by DNA helicases*. Methods, 2010. **51**(3): p. 329-335.
- Liu, Y.L. and S.C. West, *Timeline - Happy Hollidays: 40th anniversary of the Holliday junction*. Nature Reviews Molecular Cell Biology, 2004. **5**(11): p. 937-U21.
- Pannunzio, N.R., G. Watanabe, and M.R. Lieber, *Nonhomologous DNA End Joining for Repair of DNA Double-Strand Breaks*. Journal of Biological Chemistry, 2017.
- Rodríguez, A. and A. D'Andrea, *Fanconi anemia pathway*. Current Biology, 2017. **27**(18): p. R986-R988.
- Knies, K., et al., *Biallelic mutations in the ubiquitin ligase RFW3 cause Fanconi anemia*. Journal of Clinical Investigation, 2017. **127**(8): p. 3013-3027.
- Bogliolo, M. and J. Surrallés, *Fanconi anemia: a model disease for studies on human genetics and advanced therapeutics*. Current Opinion in Genetics & Development, 2015. **33**: p. 32-40.
- Nagaraju, G., et al., *Differential regulation of short- and long-tract gene conversion between sister chromatids by Rad51C*. Molecular and Cellular Biology, 2006. **26**(21): p. 8075-8086.

Agradecimientos

CONACYT proyecto SEP-CONACYT 243102, PAPIIT, proyecto IA202615, Instituto Nacional de Pediatría, proyecto 41/2014.

38. Mitelman, F., B. Johansson, and F. Mertens, *The impact of translocations and gene fusions on cancer causation*. Nature Reviews Cancer, 2007. **7**(4): p. 233-245.
39. Han, L. and K. Yu, *Altered kinetics of nonhomologous end joining and class switch recombination in ligase IV-deficient B cells*. The Journal of Experimental Medicine, 2008. **205**(12): p. 2745-2753.
40. Ambrose, M. and R.A. Gatti, *Pathogenesis of ataxia-telangiectasia: the next generation of ATM functions*. Blood, 2013. **121**(20): p. 4036-4045.
41. Frías, S., et al., *Síndromes de Inestabilidad Genómica.*, in *Genética Médica*. 2018: Ciudad de México.
42. Teive, H.A.G., et al., *Ataxia-telangiectasia - A historical review and a proposal for a new designation: ATM syndrome*. Journal of the Neurological Sciences, 2015. **355**(1-2): p. 3-6.
43. Chrzanoska, K.H., et al., *Nijmegen breakage syndrome (NBS)*. Orphanet Journal of Rare Diseases, 2012. **7**.
44. Arora, H., et al., *Bloom syndrome*. International Journal of Dermatology, 2014. **53**(7): p. 798-802.
45. Owen, N., et al., *Bloom Syndrome Radials Are Predominantly Non-Homologous and Are Suppressed by Phosphorylated BLM*. Cytogenetic and Genome Research, 2014. **144**(4): p. 255-263.
46. García-de Teresa, B., et al., *Diagnóstico clínico y de laboratorio de la anemia de Fanconi*. 33, 2012. **1**(Acta Pediátrica de México): p. 6.
47. García de Teresa, B., A. Rodríguez, and S. Frías, *Estudio multidisciplinario del paciente con anemia de Fanconi*. Acta pediátrica de México, 2016. **37**: p. 54-59.
48. Esmer, C., et al., *DEB test for Fanconi anemia detection in patients with atypical phenotypes*. American Journal of Medical Genetics Part A, 2004. **124A**(1): p. 35-39.
49. Alter, B.P. and N. Giri, *Thinking of VACTERL-H? Rule Out Fanconi Anemia According to PHENOS*. American Journal of Medical Genetics Part A, 2016. **170**(6): p. 1520-1524.



DRA. SARA FRÍAS

Bióloga, egresada de la Facultad de Ciencias, UNAM. Maestría y Doctorado en Ciencias Biológicas, UNAM. Cuenta con un posdoctorado en el National Institute of Environmental Health/Lawrence Livermore National Laboratory,

California EEUU.

Investigadora Titular, Coordinadora de la Unidad Periférica en el Instituto Nacional de Pediatría, del Instituto de Investigaciones Biomédicas de la UNAM /Investigadora en Ciencias Médicas, Laboratorio de Citogenética INP. Profesora definitiva, Biología Molecular de la Célula III, Facultad de Ciencias UNAM.

Es tutora de diversos programas de posgrado. Ha publicado 73 artículos nacionales e internacionales, 13 capítulos de libro, con más de 700 citas.

Es miembro del Sistema Nacional de Investigadores nivel 2. Ha estado involucrada en la formación de más de 40 recursos humanos desde licenciatura hasta Doctorado. Sus líneas de investigación son el a) Estudio del Fenotipo clínico, celular y molecular de la Anemia de Fanconi b) Genética y Cáncer c) Origen y etiología de las aberraciones cromosómicas.



Article

Fanconi Anemia Patients from an Indigenous Community in Mexico Carry a New Founder Pathogenic Variant in *FANCG*

Pedro Reyes ^{1,2,3} , Benilde García-de Teresa ¹ , Ulises Juárez ¹ , Fernando Pérez-Villatoro ¹ ,
Moisés O. Fiesco-Roa ^{1,4} , Alfredo Rodríguez ^{2,5} , Bertha Molina ¹ , María Teresa Villarreal-Molina ⁶ ,
Jorge Meléndez-Zajgla ⁷ , Alessandra Carnevale ⁸ , Leda Torres ^{1,*} and Sara Frias ^{1,2,*}

¹ Laboratorio de Citogenética, Instituto Nacional de Pediatría, Ciudad de México 04530, Mexico; reyesj87@gmail.com (P.R.); b.garciadeteresa@gmail.com (B.G.-d.T.); ehatlujf@gmail.com (U.J.); frpvillatoro@gmail.com (F.P.-V.); fiescoroa@facmed.unam.mx (M.O.F.-R.); berthamolina@yahoo.com.mx (B.M.)

² Departamento de Medicina Genómica y Toxicología Ambiental, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Apartado Postal 70228, Ciudad de México 04510, Mexico; alfredo.rodriguez@iibiomedicas.unam.mx

³ Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México, Ciudad de México 04510, Mexico

⁴ Maestría y Doctorado en Ciencias Médicas, Odontológicas y de la Salud, Universidad Nacional Autónoma de México, Ciudad de México 04510, Mexico

⁵ Instituto Nacional de Pediatría, Ciudad de México 04530, Mexico

⁶ Laboratorio de Genómica de Enfermedades Cardiovasculares, Instituto Nacional de Medicina Genómica, Ciudad de México 14610, Mexico; mvillareal@inmegen.gob.mx

⁷ Laboratorio de Genómica Funcional del Cáncer, Instituto Nacional de Medicina Genómica, Ciudad de México 14610, Mexico; jmelendez@inmegen.gob.mx

⁸ Laboratorio de Enfermedades Mendelianas, Instituto Nacional de Medicina Genómica, Ciudad de México 14610, Mexico; acarnevale@inmegen.gob.mx

* Correspondence: ledatorres@gmail.com (L.T.); sarafrias@iibiomedicas.unam.mx (S.F.)



Citation: Reyes, P.; García-de Teresa, B.; Juárez, U.; Pérez-Villatoro, F.; Fiesco-Roa, M.O.; Rodríguez, A.; Molina, B.; Villarreal-Molina, M.T.; Meléndez-Zajgla, J.; Carnevale, A.; et al. Fanconi Anemia Patients from an Indigenous Community in Mexico Carry a New Founder Pathogenic Variant in *FANCG*. *Int. J. Mol. Sci.* **2022**, *23*, 2334. <https://doi.org/10.3390/ijms23042334>

Academic Editor: Kirk McManus

Received: 24 December 2021

Accepted: 8 January 2022

Published: 20 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Fanconi anemia (FA) is a rare genetic disorder caused by pathogenic variants (PV) in at least 22 genes, which cooperate in the Fanconi anemia/Breast Cancer (FA/BRCA) pathway to maintain genome stability. PV in *FANCA*, *FANCC*, and *FANCG* account for most cases (~90%). This study evaluated the chromosomal, molecular, and physical phenotypic findings of a novel founder *FANCG* PV, identified in three patients with FA from the *Mixe* community of Oaxaca, Mexico. All patients presented chromosomal instability and a homozygous PV, *FANCG*: c.511-3_511-2delCA, identified by next-generation sequencing analysis. Bioinformatic predictions suggest that this deletion disrupts a splice acceptor site promoting the exon 5 skipping. Analysis of Cytoscan 750 K arrays for haplotyping and global ancestry supported the Mexican origin and founder effect of the variant, reaffirming the high frequency of founder PV in *FANCG*. The degree of bone marrow failure and physical findings (described through the acronyms VACTERL-H and PHENOS) were used to depict the phenotype of the patients. Despite having a similar frequency of chromosomal aberrations and genetic constitution, the phenotype showed a wide spectrum of severity. The identification of a founder PV could help for a systematic and accurate genetic screening of patients with FA suspicion in this population.

Keywords: Fanconi anemia; chromosome instability; *FANCG*; splicing; founder pathogenic variant; *Mixe* indigenous group

1. Introduction

Fanconi anemia (FA) is a rare genetic disorder characterized by chromosomal instability, a high predisposition to physical developmental abnormalities, progressive bone marrow failure, solid tumors, and hematological malignancies [1,2]. Genetic heterogeneity is a striking feature of FA. Germline pathogenic variants (PV) in 22 FANCG genes (*FANCA*

to FANCW) have been so far associated with the FA phenotype. The protein products of the *FANCA* genes participate in the Fanconi anemia/Breast Cancer (FA/BRCA) pathway, a biochemical network that regulates DNA damage repair in response to DNA interstrand crosslinks, maintains genomic stability during DNA replication, and assists other cellular processes [3–5].

A large number of genes and hundreds of unique PV have been associated with the FA phenotype, making allelic and locus heterogeneity the rule. Worldwide, the most frequent disease-causing PV among patients with FA occur in *FANCA* (~64%), *FANCC* (~12%), and *FANCG* (~8%) [6,7]. Recurrent PV have been identified in specific ethnic backgrounds due to a founder effect [8]. Notable examples, explained by high rates of carrier individuals, include the *FANCA* c.295C>T in Spanish Gypsies [9], *FANCC* c.456+4A>T in Ashkenazi Jews [10], and *FANCC* c.67delG in the Mennonite Community [11,12].

Regarding *FANCG* (MIM#602956), founder PV have been reported in Portuguese-Brazilian (c.1077-2A>G), Korean-Japanese (c.307+1G>C), French-Acadian (c.1480+1G>C), and Black South African (c.637_643del) populations [13,14]. *FANCG* gene covers around 6 kilobases (kb) that are mapped to chromosome 9p13; it comprises 14 exons and encodes a 2.6 kb messenger RNA (mRNA) transcript that is translated to a 622 amino acid protein with a molecular weight of 68 kilodaltons [15], this protein has seven tetratricopeptide repeat (TPR) motifs, which are required to mediate its protein-protein interactions [15,16]. *FANCG* protein is an integral component of the FA-core complex, which is a dynamic E3 ubiquitin ligase protein assembly that catalyzes the mono-ubiquitination of the FANCD2-FANCI heterodimer. This mono-ubiquitination is a critical event in the recruitment of the repair machinery to the DNA lesion as part of the FA/BRCA pathway activation [17–19].

Here, we report the novel PV in *FANCG*:c.511-3_511-2delCA, detected in three patients from the *Mixe* indigenous community in Mexico. We used a comprehensive molecular and clinical approach to elucidate the consequences of this PV and we provide evidence for a founder effect.

2. Results

2.1. Chromosomal Breakage Analysis Confirmed a FA Diagnosis

Chromosomal breakage analysis with diepoxybutane (DEB) was performed in the three patients to support the suspicion of a clinical FA diagnosis. The number of spontaneous and DEB-induced aberrations, as well as the presence of radial chromosomal figures, were considered criteria to establish the diagnosis of FA. In all three cases, the frequency of radial exchange figures and chromatid breaks showed an increase in DEB-induced samples, compared to those recorded in the spontaneous aberration cultures. Hypersensitivity to DEB treatment in patients was similar to that observed in the FA-positive control (VU817 cell line), while induced aberration frequency observed in the patients was higher than the historical average reported in our laboratory for healthy individuals (negative controls). These results confirmed the FA diagnosis in the patients, who were identified as FANCC32, FANCC143, and FANCC155 (Figure 1).

2.2. Pathogenic Variant Found in *FANCG*: c.511-3_511-2delCA Is Predicted to Induce Exon 5 Skipping

Targeted massively parallel DNA sequencing from the three patients with FA revealed a homozygous two-base deletion on *FANCG* (Supplementary Figure S1A and Table 1). This alteration is located 2 base pair (bp) downstream from the nearest acceptor splice site between intron 4 and exon 5. The genotype was subsequently validated via Sanger sequencing (Supplementary Figure S1B). According to gnomAD (<https://gnomad.broadinstitute.org/>, accessed on 6 December 2021) this variant (rs1491369358) has only been observed in 1/251,480 alleles, corresponding to a worldwide frequency of 0.00000398. Further analysis of the effect of this novel variant, following the recommended American College of Medical Genetics (ACMG) criteria, supports its pathogenicity.

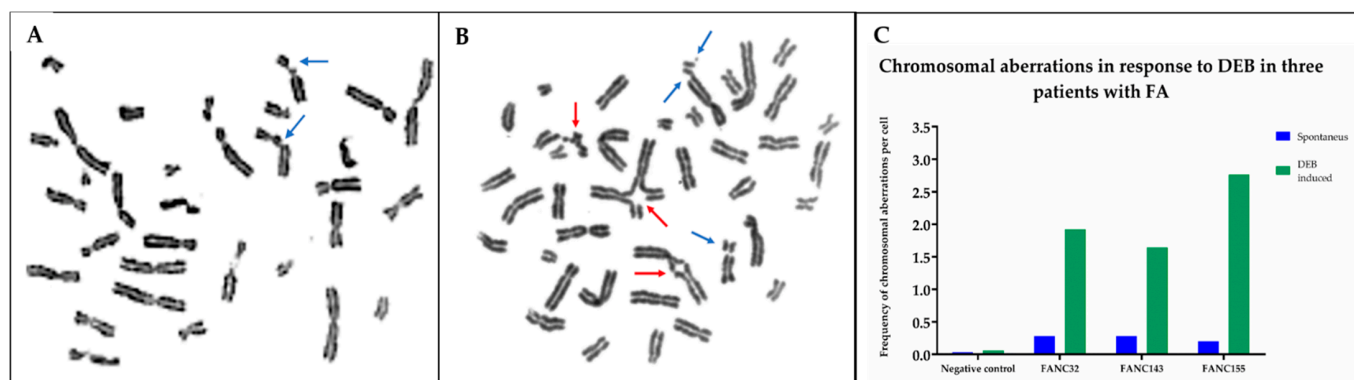


Figure 1. DEB-induced chromosomal aberrations in lymphocytes from healthy and FA individuals. DEB: diepoxybutane. (A) Representative metaphase from a healthy individual (negative control); chromosome aberrations are mainly chromatid breaks. (B) Representative metaphase from one of the FA patients; increased chromosome aberrations are observed showing chromatid breaks and radial figures, characteristics of patients with FA. Blue arrows show chromatid and isochromatid breaks and red arrows show radial figures. (C) Frequency of chromosomal aberrations per cell for each FA patient analyzed in this study.

Table 1. Features of the pathogenic variant found in *FANCG*.

Reference Sequence	DNA Change (Genomic, hg19)	ACMG Criteria Variant Classification	HGVS Nomenclature
NG_007312.1	g.35077398_35077399del	PVS1, PM2, PP3	<i>FANCG</i> (NM_004629.2): c.511-3_511-2delCA

Abbreviations: ACMG: American College of Medical Genetics; c.511-3_511-2delCA: two bp deletion in coding DNA; g.35077398_35077399del: two bp deletion in genomic DNA; hg19: Homo sapiens (human) genome assembly GRCh37; HGVS: Human Genome Variation Society; PM2: Pathogenic Moderate 2; PP3: Pathogenic Supporting 3; PVS1: Pathogenic very strong 1. For a complete description of these criteria refer to [20].

Inspection of the genomic region where the variant was found suggests that this change could disrupt mRNA processing. In silico analysis using The Human Splice Finder tool revealed that *FANCG*:c.511-3_511-2delCA impairs the wild-type acceptor site, most probably affecting splicing. MaxEntScan rated this PV with a score of 1.1, with respect to a reference score of 9.63, predicting a deleterious modification of this splice site. Finally, CRYP—SKYP tool indicated that the deletion may promote exon 5 skipping (probability of cryptic site activation = 0.40). These results consensually predict that this PV in the acceptor splice site leads to loss of *FANCG* exon 5. The consequence of exon 5 skipping would lead to reading frame shift, and generation of a premature stop codon at position p.S171Vfs5*, resulting in a truncated *FANCG* protein (Supplementary Figure S1C).

2.3. Reduced Genetic Variation in Locus g.29754068-38771831 Supporting a Founder Effect for *FANCG*:c.511-3_511-2delCA in the Mixe Population

The three patients with FA described in this study come from an isolated mountainous region in the state of Oaxaca (southern Mexico). They self-identified as belonging to the *Mixe* indigenous group and shared a surname. Although a kin relationship was unknown, we investigated the possibility of a common haplotype using high-density microarrays (CytoScan 750K array). Several long-contiguous stretches of homozygosity (LCSH) > 3 Mb were shared in autosomal regions by these three patients with FA (Table 2). Some of these LCSH have previously been reported in the HapMap and the 1000 Genomes Project (1KGP) in populations of Asian (CHB), Northern European (CEU) and American (AMR) origin (Supplementary Table S1) [21–24]. We found two shared zones among the LCSH regions not previously reported. The first one in chromosome 5q spanned 2935 kb (g. 101035497-98100800) and the other one in chromosome 9p spanned 9018 kb (g.29754068-38771831); the

latter, importantly, included the *FANCG* locus. Inbreeding coefficient (F), calculated with the total LCSH extension and used to estimate the degree of kinship, suggested a possible parental relationship between the three patients (Table 2).

Table 2. Native component, homozygosity, and coefficient of inbreeding in the three *Mixe* patients with Fanconi anemia.

Patient ID	Total Autosomal LCSH > 3 Mb (kb)	Percentage of Native Component	Autosomal Homozygosity (%)	F	Probable Parental Relationship
FANC32	140214361	79.17%	4.9%	0.0487	Fourth-degree
FANC143	203913825	85.64%	7.1%	0.0708	Third-degree
FANC155	249262167	88.53%	8.7%	0.0865	Third-degree

Abbreviations: F: Inbreeding coefficient; LCSH: Long-contiguous stretches of homozygosity.

When we analyzed the g.29754068-38771831 region, we discovered that 439 out of 544 single nucleotide variants (SNV) (80.7%) included in the CytoScan 750K array were common among the three patients. We compared the allelic frequency of these SNVs with the frequency reported in the 1KGP for the global and the Mexican population (MXL). We found that 99.32% of the alleles of our patients are coincident frequencies, compared to the global frequencies and the MXL group, whereas the remaining 0.68% has a different frequency. These results confirm that patients share a common haplotype of 30 SNVs around *FANCG*, supporting the discovery of a founder effect (Supplementary Table S2).

Finally, to ascertain that our patients share the Mexican population (MXL) genetic background, as reported in the 1KGP, a global ancestry analysis was performed. Considering that Mexicans present a remarkable genetic diversity due to population admixture, we decided to contrast all SNVs data of these 3 *Mixe* patients against other populations beside the MXL. These populations included the following: Native-Americans (NAT), Utah residents with Northern and Western European ancestry (CEU), and Yoruba in Ibadan, Nigeria (YRI). Through similar approach it has been demonstrated that the MXL population presents an ancestry pattern with Native and European components predominance (NAT: 46.59% and CEU: 48.44%, respectively) [25,26]. Our results indicated that the patients have a proportion of the NAT as high as 79 to 88% (Table 2), with lower CEU and YRI contributions, when compared to the MXL population. Although diversity is observed, the predominance of CEU and NAT ancestry pattern is conserved in both MXL and *Mixe* patients (Figure 2).

2.4. The Fraction of Pathogenic Variants Linked to a Founder Effect in *FANCG* Is the Highest among the Most Common *FANC* Genes

Most founder variants described so far in FA are clustered together in the three most frequently mutated *FANC* genes (i.e., *FANCA*, *FANCC*, and *FANCG*). The PV described here adds a new founder PV (FPV) in *FANCG* (Table 3, Figure 3). The literature analysis indicates that the number of FPV, validated through haplotype analysis, was eleven, three, and eight for *FANCA*, *FANCC*, *FANCG*, respectively; this study adds the ninth FPV in *FANCG* gene. *FANCG* has the lowest number of unique variants among the most frequently affected genes in FA. There were differences between the proportion of founder and non-founder variants among these genes (Figure 4). The Fisher's exact test to compare the proportion of FPV between *FANCG* and *FANCA* was $p < 0.0001$, whereas between *FANCG* and *FANCC* was $p < 0.0001$ (Figure 4).

2.5. Evaluation of the Phenotype of *Mixe* Patients with FA in the Framework of the Reported *FANCG* Pathogenic Variant

Medical records of each patient were analyzed to identify physical findings, mainly those described in VACTERL-H (Vertebral, Anal, Cardiac, Tracheo-esophageal fistula, Esophageal atresia, Renal, upper Limb, and Hydrocephalus) association (MIM#192350) and PHENOS (skin Pigmentation, small Head, small Eyes, Nervous system, Otology, Short

stature) acronyms [30]. Likewise, hematologic data analysis was used to classify the severity of the bone marrow failure (BMF).

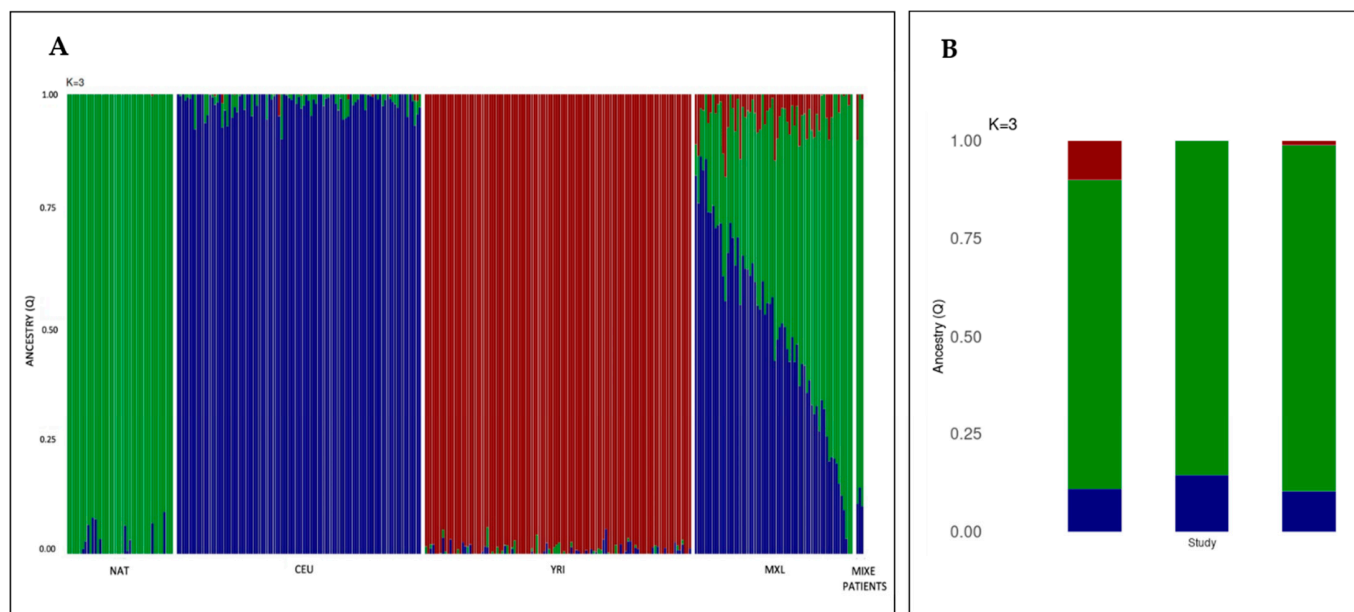


Figure 2. *Mixe* patients with FA present an ancestry pattern that have the hallmarks of the genetic diversity of the Mexican population. **(A)** Barplot showing the inference of global ancestry for the *Mixe* and Mexican individuals using ADMIXTURE, population reference panels include NAT, CEU, and YRI (K = 3 model). Each individual is depicted as a vertical bar. Colors represent the percentage of ancestry assigned to each cluster for each individual. Green, blue, and red colors indicate Native-Americans (NAT), Northern Europeans (CEU), and Yorubas (YRI) populations, respectively. The two panels located at the extreme right of the figure show the three-way ancestry components for the admixed Mexican (MXL) population and the *Mixe* patients (last three bars). **(B)** Detailed bar plot of the three *Mixe* patients with FA, left to right FANC32, FANC143 and FANC155.

Table 3. Founder pathogenic variants reported in *FANCG*.

Pathogenic Variant	Location	Effect	Geographic/ Ethnic Background	Reference
c.307+1G > C	Intron 3	Aberrant splicing	Japanese-Korean	[27]
c.313G > T (p.Glu105Ter)	Exon 4	Truncated protein, null variant	German	[28]
c.637_643delTACCGCC (p.Tyr213LysfsTer6)	Exon 5	Truncated protein, null variant	South African	[14]
c.1066C > T (p.Gln356Ter)	Exon 8	Truncated protein, null variant	Japanese	[29]
c.1077-2A > C	Intron 8	Aberrant splicing	Portuguese-Brazilian	[13]
c.1480+1G > C	Intron 11	Aberrant splicing?	French-Acadian	[13]
c.1589_1591delATA (p.Asp530_Thr531delinsAla)	Exon 12	Reduce protein activity, Hypomorphic variant	Korean	[29]
c.1649delC (p.Thr550IlefsTer9)	Exon 13	Truncated protein, null variant	Turkish	[28]
c.511-3_511-2delCA	Intron 4	Aberrant splicing and truncated protein	<i>Mixe</i> (Mexican)	Present study

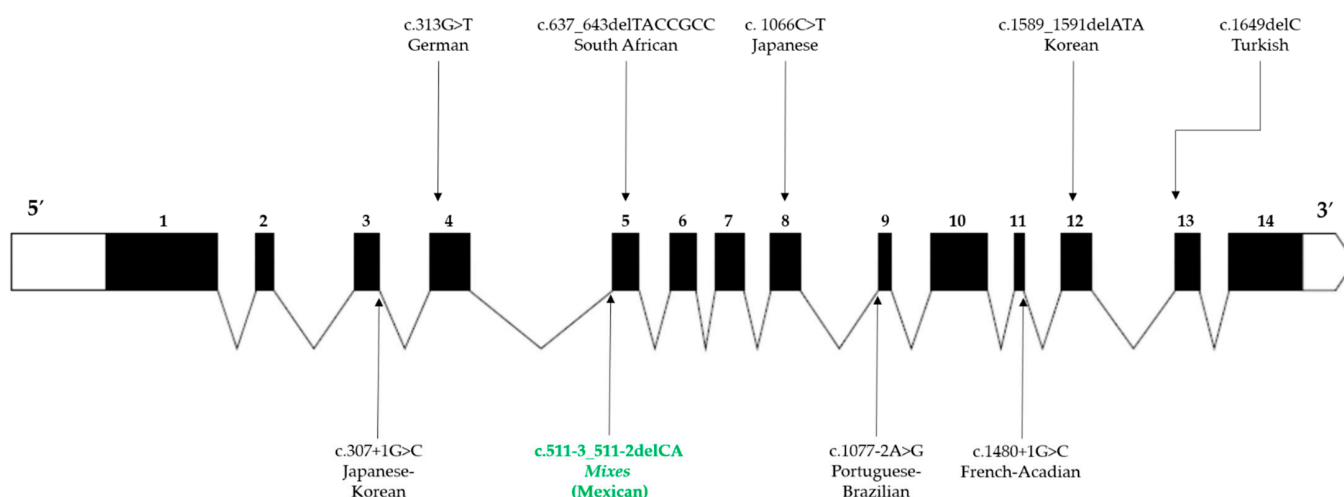


Figure 3. Worldwide founder pathogenic variants reported in *FANCG*. The structure of *FANCG* is described in 5'-3' orientation. The untranslated regions are illustrated at the ends of the gene (white areas), the exons are represented by numbered black boxes, and the angled lines correspond to the introns. The location of the variants is indicated by vertical arrows, those located at the bottom of the figure represent splicing variants, while deletions and missense variants are located at the top. The variant described in this study is represented by green letters.

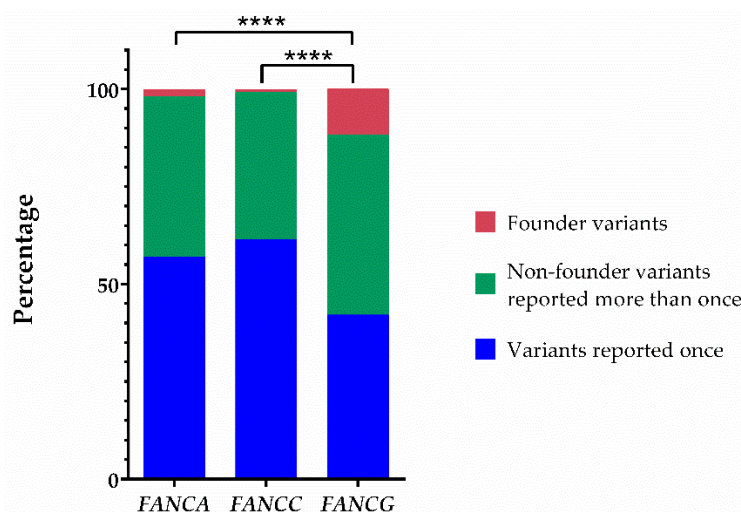


Figure 4. The proportion of variants with founder effect in *FANCG* is the highest among the most frequently reported genotypes. Proportions of founder pathogenic variants, non-founder pathogenic variants reported more than once, and pathogenic variants reported just once in *FANCA*, *FANCC*, and *FANCG*. The significance level of Fisher's exact test is represented by four asterisks ($p < 0.0001$).

FANC32 patient was the third of four children, born to non-consanguineous parents. His older sister had FA diagnosis and died due to BMF. The patient's pregnancy was uneventful; he was delivered vaginally at 40 weeks of gestation (WG). At birth his weight was 2500 g (<3rd percentile, HP:0001518) and he had hypotonia (HP:0001252). Developmental milestones were delayed (HP:0012758), beginning to walk at 36 months. He was sent to our clinic at seven years old (y.o.) due to BMF with a history of recurrent upper respiratory tract infections (HP:0002788). Bone marrow aspiration and biopsy performed at age nine y.o. reported 50% cellularity with erythroid series and megakaryocyte predominance and no signs of fibrosis. He was initially treated with steroids and transfusions when required. He later started evaluation for hematopoietic stem-cell transplantation (HSCT); unfortunately, the patient stopped attending follow-up visits. At 14 y.o. anthropometric evaluation showed low weight (−2.31 standard deviation score (SDS), HP:0004325), height (−2.73 SDS,

HP:0004322), and head circumference (-2.07 SDS, HP:0000252) for his age and sex. Physical examination revealed thenar hypoplasia of both hands (HP:0001245), generalized hyperpigmentation (HP:0000953), *café au lait* spots (HP:0000957), and melanonychia (HP:0100644). Cardiac and renal structural abnormalities were not found through sonographic evaluation.

FANC143 patient was the first child of non-consanguineous parents. Her pregnancy was complicated by fetal distress (HP:0025116) and oligohydramnios (HP:0001562) which warranted pregnancy termination by C-section at 36 WG. The birth weight was 1590 g (<3 rd percentile, HP:0001518), she remained hospitalized her first month of life. Bilateral radial ray alteration (HP:0410049) and left hip dysplasia (HP:0001385) were identified. She had developmental delay (HP:0012758), particularly affecting language (HP:0002463). The patient was referred to our clinic when she was nearly seven y.o. due to BMF. She was initially treated with transfusions when needed and steroids. Physical examination evidenced low weight (-2.15 SDS, HP:0004325), short stature (-2.90 SDS, HP:0004322), and microcephaly (-3.23 SDS, HP:0000252). She had up-slanted palpebral fissures (HP:0000582), bifid uvula (HP:0000193), bilateral microtia (HP:0008551), with folded pinnae of the right ear (HP:0000396), agenesis of the right thumb (HP:0009777), floating hypoplastic left thumb (HP:0009601), generalized skin hyperpigmentation (HP:0000953), and *café au lait* spots (HP:0000957). Tympanometry and brainstem auditory evoked potentials showed right unilateral hearing loss (HP:0000365). Spinal X-rays showed spina bifida in S1 (HP:0004614) and a lumbosacral hemivertebra (HP:0008439). Renal and cardiac malformations were ruled out with imaging methods. Brain magnetic resonance imaging (MRI) revealed cortico-subcortical atrophy (HP:0002120 and HP:0012157), supratentorial ventriculomegaly (HP:0002119), and encephalomalacia (HP:0040197). The patient underwent HSCT at 8.6 y.o. and died a month later due to graft versus host disease complications and septic shock.

FANC155 patient was the younger of two children, born to non-consanguineous parents. Her family health history showed that two great-aunts died of breast cancer in their early thirties. She had a maternal uncle with a history of infertility and vertebral defects. The mother reported an uneventful pregnancy. She was vaginally delivered after 40 weeks, with a birth weight of 2400 g (<3 rd percentile, HP:0001518). She had developmental delay (HP:0012758) with language disability (HP:0002463). At six y.o. she was evaluated at a local hospital due to a history of recurrent epistaxis (HP:0004406) and pallor (HP:0000980). After evaluation, a diagnosis of aplastic anemia (HP:0001915) was reached. She received blood transfusions when required and steroid therapy. At 12 y.o. she was referred to our clinic for HSCT. Physical examination evidenced weight (-1.26 SDS) and head circumference (-0.86 SDS) within normal range but low height (-2.18 SDS, HP:0004322) for her age and sex. She had up-slanted palpebral fissures (HP:0000582), ptosis (HP:0000508), simple myopic astigmatism (HP:0500041), hypoplastic helix in both ears (HP:0011039), left hypoplastic thenar region (HP:0001245) with limited ipsilateral thumb flexion, generalized hyperpigmentation (HP:0000953), and *café au lait* spots (HP:0000957). Audiometric findings were unremarkable. Brain MRI, renal ultrasonography, and echocardiography did not show alterations. She underwent HSCT from a non-related donor but died a couple of days after the procedure due to septic shock.

Table 4 summarizes the three patients' clinical phenotypes. The median age at FA diagnosis was 6.45 y.o. All patients had short stature, skin pigmentation changes, and radial ray anomalies. Respecting VACTERL-H abnormalities, one patient had vertebral and upper limb (radial ray) malformations, and the other two patients had only upper limb alterations; none of them met criteria for VACTERL-H association ($\geq 3/8$ features). All three patients had at least three PHENOS anomalies; however, only one had $\geq 4/6$ PHENOS features. Regarding hematologic phenotype, the three patients had moderate to severe BMF.

We compared the physical manifestations observed in the patients reported here to those described in the literature for cases with PV in FANCG, we excluded reports in which the phenotypes were not mentioned or individually described. We identified 84 cases with PV in FANCG. Remarkably, none of these had any of the 8 founder-effect PV listed in Table 3. The most frequent physical characteristics were short stature (54%), skin pigmentation

changes (38%), upper limb (radial ray) abnormalities (31%), microcephaly (29%), and renal malformations (20%). The rest of abnormalities were present in less than 20%. The type and frequency of phenotypic anomalies are depicted in Figure 5. Three out of 84 patients met criteria for VACTERL-H ($\geq 3/8$ features), seven out of 84 had $\geq 4/6$ PHENOS features, and only one patient had VACTERL-H plus PHENOS. The median age at diagnosis was 6 y.o. and the median age at report was 9 y.o. The male-to-female ratio was 1:0.64 ($p = 0.5$).

Table 4. Summary of the clinical phenotype of the patients described in this study.

Patient ID	Sex	Family History of FA	Age at Diagnosis	Hematologic Phenotype at Initial Evaluation					VACTERL-H Features	PHENOS Features
				Hb (g/dL)	MCV (fL)	ANC (cells/ μ L)	PLT (cells/ μ L)	BMF Status		
FANC32	M	Sister	6.45	11.4	100.0	700	26,000	Moderate BMF	L	PHS
FANC143 [†]	F	No	3.6	8.9	94.1	900	15,000	Moderate BMF	VL	PHNOS
FANC155 [†]	F	No	9.3	11.1	99.5	220	19,000	Severe BMF	L	POS

Abbreviations: ANC: Absolut Neutrophils Counts; BMF: Bone marrow failure; F: female; Hb: Hemoglobin; FA: Fanconi anemia; M: male; MCV: Mean Corpuscular Volume; PLT: Platelets; PHENOS: skin Pigmentation, small Head, small Eyes, Nervous system, Otology, Short stature; VACTERL-H: Vertebral, Anal, Cardiac, Tracheo-esophageal fistula, Esophageal atresia, Renal, upper Limb, and Hydrocephalus; [†] Patient deceased.

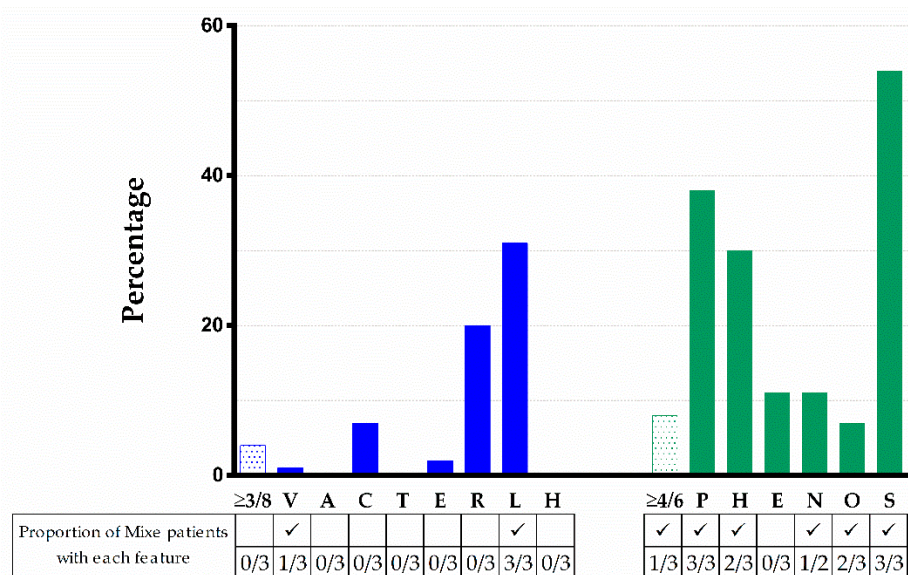


Figure 5. VACTERL-H and PHENOS features in 84 patients from the literature with a *FANCG* genotype and comparison with the three patients with FA here described. Upper limb includes abnormal thumb +/– abnormal radius. Nervous system includes structural brain malformations other than hydrocephalus. Otology comprises ear malformations and/or hearing loss. Solid blue or green bars: VACTERL-H or PHENOS; dotted bars: individual findings. Horizontal axis: abnormalities analyzed; vertical axis: percent of total cases with that abnormality. VACTERL-H association ($\geq 3/8$ features) was present in 4% and PHENOS ($\geq 4/6$ features) in 8% of patients. The proportion of *Mixe* patients with each feature is shown in the table below the figure. The character \checkmark means that this finding was found among *Mixe* patients. Imaging brain evaluation was performed in only two patients.

3. Discussion

FA is a disorder characterized by high genetic, allelic, and locus heterogeneity. The simultaneous study of several genes using next generation sequencing (NGS)-based targeted panels has advanced the recognition of variants not previously reported within known *FANCG* genes. Applying this NGS strategy, we diagnosed three patients from the same geographical region with a classic FA phenotype. We found a non-previously reported PV:

FANCG:c.511-3_511-2delCA, consisting of a two base-pair deletion affecting the splicing acceptor site of *FANCG* exon 5. This deletion is predicted to lead to an aberrant splicing site and a stop codon, generating a truncated *FANCG* protein lacking TPR motifs and affecting the function of the FA core complex. Confirmation of a founder effect was obtained by haplotype and ancestry analysis. Overall, our results represent an integrative approach to evaluate the effect of this change in *FANCG* and provide useful information to guide a potential targeted FA screening on the *Mixe* population, a well-defined ethnic group in Mexico.

Besides having the same variant, these patients also share the same geographic and ethnic origin, as well as one of their surnames. These data, in addition to the very low reported frequency of this change (rs1491369358), which has been found in a heterozygous state only in one non-FA individual, led us to explore the possibility of a founder effect for PV *FANCG*:c. 511-3_511-2delCA. The analysis of the degree of inbreeding (F value), the percentage of homozygosity and the regions with LCSH allowed us to corroborate this inference, since all the patients have large stretches of homozygosity in their genomes due to inheritance of identical ancestral genomic segments from both parents. A consequence of this genetic structure is an increased incidence of recessive diseases, like FA. Additionally, the inferred haplotype in the boundaries of the *FANCG* locus supports the hypothesis of a founder effect in this novel PV. Moreover, the high contribution of the native component of around 80% in these three patients contrasts with the average 46% found in the MXL group and supports the idea that these subjects who belong to the *Mixe* group have reduced genetic diversity (Figure 2). While only two of these patients have a native contribution of over 85%, sufficient to name them Mexican natives [31], the third one who does not reach this threshold also has the lower autosomic homozygosity and the more distant familial relationship, reflecting more mixture and illustrating the origin of mestizo-Mexicans [23].

FANCG:c.511-3_511-2delCA is the first FPV reported in Mexican patients with FA and is consistent with the demographic history of the *Mixes*; an original population that has resisted several attempts to occupy their lands since the pre-Hispanic era and the Spanish conquest. Even though changes have taken place in the political-administrative organization of the region in which they reside since then, this isolated ethnic group has preserved its language, social structure, and territory [32]. These facts also highlight their sedentary nature, which has led to limited genetic admixture [33]. Considering the appraised size of the community where they come from [34], the calculated prevalence of FA is 7.5/1,000,000, which falls in the upper rank of the estimated worldwide FA prevalence in the general population (1-9/1,000,000) [35]. Of note, FPV constitute more than 10% of the total PV discovered so far in *FANCG*, the highest proportion among the most prevalent *FANC* genes (Figure 4). Although most of the PV in FA are concentrated in *FANCA*, *FANCC*, and *FANCG*, a study suggests that they do not appear to be intrinsically more mutable than the other *FANC* genes [36]; therefore, another explanation should be sought to understand the higher prevalence of PV in these genes. On the other hand, it is unlikely that the rate of FPV in *FANCG* could confer any advantage to carriers, since half of them generate a truncated protein and others a dysfunctional one, so it is likely that they may occur due to chance. Certainly, this observation deserves further investigation.

The patients presented here have a classic FA phenotype, consisting of BMF and physical abnormalities [37]. The acronyms VACTERL-H and PHENOS include the most frequently reported physical abnormalities in patients with FA [30,38]. Among these 14 features, only five are present in over 25% of patients: short stature (43%), radial ray defects (40%), changes in skin pigmentation (37%), microcephaly (27%), and renal malformations (27%) [38]. All three patients described here share height, radial, and skin abnormalities; two also had microcephaly and ultrasonographic evaluations ruled out renal malformations.

Figure 5 shows that 4% of patients reported in the literature with PV in *FANCG* met criteria for VACTERL-H; if we consider all genotypes, this percentage rises to 12% [38]. The percentage of patients reported in the literature with PV in *FANCG* who had $\geq 4/6$ PHENOS features was similar considering all genotypes (8.3 vs. 9%). None of our patients had

VACTERL-H and only one had $\geq 4/6$ PHENOS features. None of our patients had evident small eyes, but we did not measure palpebral fissure length, and ocular imaging was not performed. Yet, we did find other ophthalmologic manifestations such as ptosis, myopia, and astigmatism (FANC155). Small eyes are described in 11% of patients with FA regardless of genotype [38], this feature stands for the “E” in the PHENOS acronym. However, the ocular phenotype in FA is wider as recently demonstrated [39] and broadening the features that conform the “E” in PHENOS should be considered.

As expected, although these three patients share the same pathogenic *FANCG* genotype and a similar genetic background, their phenotypic presentation was not homogeneous, as evidenced in other cases [40]. Yet, they illustrate that clinical presentation in FA has a broad spectrum of severity that in this case goes from a florid presentation in FANC143, which includes five PHENOS features and uncommon findings like bifid uvula, to a more discreet presentation in patient FANC155 in whom her unilateral radial abnormality was only recorded after intentional evaluation by a dysmorphologist (Table 4). Concerning the hematologic phenotype, patients with *FANCG* PV have been found to have more severe cytopenia as well as a higher frequency and an earlier diagnosis of acute myeloid leukemia or myelodysplastic syndrome than patients with *FANCA* or *FANCC* genotypes [41]. All three patients have a hematologic phenotype that merited HSCT, yet platelets and neutrophils appeared to be more severely affected than red cells. Oncologic manifestations were not found in these patients, but this observation is certainly limited not only because of the small number of individuals but also due to their young age.

To our knowledge, the only available phenotypic description of a patient cohort with a FPV *FANCG* genotype refers to Black South African individuals, with the homozygous variant c.637_643del [14]. The analysis of 35 patients with this FPV could not definitely establish a particular phenotype for this genotype, yet it was recognized that they had a high frequency of skin pigmentation alterations (97%) and that abnormalities of the upper limbs were frequent (>70%) and subtle (without radial hypoplasia). Also, the frequency of renal abnormalities (37%) was higher than reported elsewhere [42]. A second study that included 24 patients and focused in endocrinologic data, showed that 33%, 46%, and 42% had microcephaly, short stature, and low weight for age, respectively [43]. When comparing the phenotype of the three *Mixe* patients to the Black South African patients, in whom the genotype also leads to a truncated protein (p.Tyr213Lysfs*6), the anthropometric alterations appear to be less severe in the South African patients [43]. We take these observations with caution, both because of the small size of the *Mixe* cohort and the influences of environmental factors and additional *loci* in these multifactorial traits cannot be overlooked. Therefore, a study in this geographical region should be warranted.

In conclusion, in this study we present the ninth FPV in the *FANCG* gene, thereby contributing to reaffirm that this gene has the highest proportion of FPV among the most frequently affected *FANC* genes. Despite having similar genetic and environmental backgrounds, the phenotype of these three *Mixe* patients agrees with the highly variable dysmorphological phenotype of patients with PV in *FANCG* described in the literature. The identification of the founder *FANCG*:c.511-3_511-2delCA PV in these patients constitutes a fundamental step for the systematic and accurate genetic screening of patients with FA in the *Mixe* indigenous group. This could lead to carrier identification, appropriate genetic counseling for families with FA, and optimization of the search for potential bone marrow transplant donors. Finally, the identification of a PV in a population with a high prevalence of the disease represents an opportunity to study the phenotypic effect of the variant in a population with a homogeneous genetic background.

4. Materials and Methods

4.1. Editorial Policies and Ethical Considerations

This study was approved by the Ethics and Research Committees of the Instituto Nacional de Pediatría (INP) in Mexico City. All patients provided written informed consent

to participate in this study (INP 041-2014) and blood samples of each participant were collected by peripheral venipuncture.

4.2. Patients

Three patients with cytopenias were independently referred for a medical evaluation to the INP. These patients came from the same geographic region in the southern part of Mexico (Oaxaca State). All patients underwent a comprehensive physical examination by trained medical geneticists. Clinical data including age, sex, parental consanguinity, family history, and anthropometric measurements (weight, height, and head circumference) were recorded. Hematological information, as full blood counts at presentation and subsequent monitoring, bone marrow biopsy results, age of development of bone marrow failure, and transfusions were documented through retrospective review of each patient's medical file. Available imaging studies were evaluated. For the phenotype description, the Human Phenotype Ontology was used [44].

4.3. Chromosomal Breakage Test

A diepoxybutane (DEB) test was performed in all patients. Two-paired lymphocyte cultures per blood sample were cultured in RPMI-1640 medium (Gibco, BRL, Grand Island, NY, USA) and stimulated with phytohemagglutinin (Gibco, BRL, Grand Island, NY, USA). To induce DNA damage, 0.1 µg/mL of DEB (Sigma, St Louis, MO, USA) was added from the start of the incubation to half of the cultures; the untreated cultures were used for the analysis of spontaneous chromosomal breakage. In parallel, cultures of blood samples from healthy individuals as negative controls and the FANCA cell line FA VU817 as a positive control were performed. Harvesting of all cultures was carried out after of 72 h at 37 °C in a 5% CO₂ incubator. Metaphase spreads and staining were performed according to standard protocols [45]. We evaluated the frequency of chromosome aberrations including breaks, fragments, dicentrics, rings, and radial figures. The frequency of aberrations per cell, as well as the percentage of aberrant cells, were calculated.

4.4. Genotyping

4.4.1. Genomic DNA Extraction

The Gentra Puregene Kit (Qiagen, Venlo, Limburg, NL, USA) was used for the extraction of total genomic DNA from peripheral blood samples. The concentrations of the DNA were determined on a Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA) and DNA integrity was verified by agarose gel electrophoresis.

4.4.2. Targeted Next-Generation Sequencing

A customized HaloPlex panel (Agilent Technologies, Santa Clara, CA, USA) was designed to target 16 FA genes (*FANCA*—*A, B, C, D1, D2, E, F, G, I, J, L, M, N, O, P* and *Q*). The final design included all coding exons and 50 bp of their flanking 5' and 3' intronic regions, with a coverage of 99.5%. Agilent's HaloPlex Enrichment System protocol was followed for library preparation and paired-end sequencing was performed in the MiSeq System platform (Illumina, San Diego, CA, USA). Raw reads were filtered and mapped to the GRCh37/hg19 reference genome with the Burrows-Wheeler Aligner (BWA). The overall mean sequencing depth of the samples was 1000X. Alignment and variant calling were performed using the Genome Analysis Toolkit (GATK), version 4.0.3. A pipeline was designed to maximize the accuracy of variant calls, according to GATK Best Practices recommendations [46,47]. Variants were described following the guidelines proposed by the HGVS nomenclature [48]. The classification of variants, according to the five-tier criteria of the ACMG guidelines [20] as well as the estimated global frequency, were assessed using the online tool VarSome [49].

4.4.3. Sanger Sequencing

FANCG variant-specific primers were designed to encompass the splicing acceptor site between intron 4 and exon 5, using the Primer3 software (<http://bioinfo.ut.ee/primer3-0.4.0/>, accessed on 22 August 2019). The selected primers were as follows: forward 5'- GACCTTG-GCGGTAGGCAAA; and reverse 5'- ATTGGGGGAAACTACAGGCA. PCR amplification was carried out from 100 ng of DNA template, according to standard protocols [50]. PCR products were purified with QIAquick kit (QIAGEN, Venlo, Limburg, NL, USA) according to manufacturer instructions. Purified amplicons were bidirectionally sequenced using the Big Dye Terminator sequencing kit and resolved on an Applied Biosystems 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The electropherograms were analyzed with Chromas V2.6.6 software (www.technelysium.com.au, accessed on 24 January 2020) and target sequences were compared to the corresponding reference sequences *FANCG* (NG_007312.1) from GenBank (<https://www.ncbi.nlm.nih.gov>, accessed on 24 January 2020).

4.5. In Silico Splice Site Analysis

Human Splicing Finder (HSF) (<http://www.umd.be/HSF/>, accessed on 5 February 2020), MaxEntScan (http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html, accessed on 5 February 2020) and CRYP-SKIP (<https://cryp-skip.img.cas.cz/>, accessed on 7 February 2020) were the three bioinformatics tools used to predict the effect of the variant on splicing signals.

4.6. High-Resolution Microarray Analysis

We performed the analysis of the gDNA with Affymetrix CytoScan 750K Arrays (Affymetrix, Santa Clara, CA, USA) on the three patients. This array provides genome-wide coverage, including 550,000 markers for detecting copy number variation and 200,436 SNP probes. We performed the procedures for DNA digestion, ligation, PCR amplification, fragmentation, labelling, denaturing and hybridization into the array according to the protocols and QC guidelines provided by the supplier. Arrays were then stained and washed in the affymetrix GeneChip Fluidics Station 450 and scanned using an Affymetrix GeneChip Scanner 3000 7G (Affymetrix, Santa Clara, CA, USA); we analyzed the files obtained with the appropriate bioinformatics tools.

4.6.1. Analysis of Long-Contiguous Stretch of Homozygosity

We visualized the LCSH in the software Chromosome Analysis Suite (ChAS) software version 4.1, provided by Affymetrix (Affymetrix, Santa Clara, CA, USA). For the analysis, we used the NetAffx 33 hg19 annotation files (<https://www.affymetrix.com/analysis/index.affx>, accessed on 20 January 2021). For LCSH > 3Mb, the analysis configuration was set at LOH with marker count = 50 and size = 3000 kb, and for LCSH > 5Mb was set at marker count = 50 and size = 5000 kb.

4.6.2. Estimation of Coefficient of Inbreeding

Individual inbreeding coefficients (F) were estimated using LCSH > 3 Mb data; F was the total length of autosomal LCSH in kb divided by the total autosomal size covered by the Cytoscan 750 array (2,881,033,286 kb for hg19). The F value of 0.25 could reflect a first-degree parental relationship, 0.125 a second-degree, 0.0625 a third-degree and 0.03125 a fourth-degree [51].

4.6.3. Haplotype Inference

We used the LSCH > 5 Mb data to perform a haplotype inference approach, specifically the LCSH region shared by the three patients that include *FANCG*. The number of SNVs identical in the three patients was calculated, and the allele frequency was reviewed to infer the haplotype. The SNVs was compared with those reported in global population and the unrelated Mexican individuals (86 samples, MXL population) found in the 1000 Genomes Project (phase 3) database (1KGP) (<http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/>

[phase3/data](#), accessed on 16 December 2021), using the LDhap tool (<https://ldlink.nci.nih.gov/?tab=ldhap>, accessed on 16 December 2021).

4.7. Ancestry Analysis

For the ancestry study, CEL files derived from the arrays were analyzed with the ChAS software to obtain the SNV genotyping calls, generated from the use of the BRLMM-P-plus algorithm. Subsequently, these results were contrasted with genotype information derived from Affymetrix high-density arrays present in the 1KGP. PLINK 1.9 software (<https://zzz.bwh.harvard.edu/plink/download.shtml/>, accessed on 25 January 2021) was used to filter and map the SNVs. SNVs with discordant strings or genotypes were corrected or removed. Finally, the global ancestry of the MXL population and the three patients was deduced through a supervised maximum likelihood ADMIXTURE approach, from $K = 2$ to $K = 3$ ancestral components. For this analysis, genotypic frequencies of Northern European (CEU), Yoruba (YRI), and Native American (NAT), reported in the 1KGP, were considered as reference parental populations.

4.8. Fanconi Anemia Variant Database Analysis

We searched the Fanconi anemia Mutation Database (displayed using Leiden Open Variation Database [LOVD, v.3.0], <https://www2.rockefeller.edu/fanconi/genes/>, accessed on 30 November 2021) for *FANCA*, *FANCC*, and *FANCG*. We first categorized the variant according to the number of times it has been reported. In those where five or more patients were acknowledged, we reviewed the cited publications in order to establish if a haplotype analysis demonstrated a founder effect. We complemented this with a PubMed/MEDLINE search using the terms: “*FANCA* OR *FANCC* OR *FANCG*” AND “mutation OR variant” AND “founder OR haplotype”. We classified the variants as founder PV (FPV), when demonstrated by haplotype analysis, repeatedly reported non-founder PV (recurrent, but not demonstrated as FPV by haplotype analysis), and those reported just once. We compared the proportion of these variant categories between *FANCG* and the other two genes *FANCA* and *FANCC*. Fisher’s exact test (p -values < 0.05) was used to statistically test these differences.

4.9. Phenotype Analysis

We searched PubMed/MEDLINE for publications limited to human subjects through 1 December 2020, using the terms “Fanconi” AND “anemia”. We collected and analyzed all cases with a *FANCG* genotype in whom individual patient’s phenotypes were detailed. We described frequencies of VACTERL-H (Vertebral, Anal, Cardiac, Tracheo-esophageal fistula, Esophageal atresia, Renal, upper Limb, and Hydrocephalus) and PHENOS (skin Pigmentation, small Head, small Eyes, Nervous system, Otology, Short stature) features. VACTERL-H association was considered if the patient had $\geq 3/8$ features, and for PHENOS $\geq 4/6$ features were needed. We extracted clinical information from publications reporting in *FANCG* FPV (as described in Section 4.8) to compare it to the phenotype of the patients reported here. The statistical analyses were performed using Microsoft Excel Office 365 (Microsoft, Redmond, WA, USA) and R (R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>, accessed on 16 December 2021. We used Fisher’s exact test and p -values < 0.05 were significant.

Supplementary Materials: Supplementary Materials can be found at: <https://www.mdpi.com/article/10.3390/ijms23042334/s1>.

Author Contributions: Conceptualization, B.G.-d.T., L.T. and S.F.; Data curation, B.G.-d.T., U.J., M.O.F.-R., L.T. and S.F.; Formal analysis, P.R., B.G.-d.T., U.J., F.P.-V., M.O.F.-R., A.R., B.M., M.T.V.-M., J.M.-Z., L.T. and S.F.; Funding acquisition, A.C., L.T. and S.F.; Investigation, S.F.; Methodology, P.R., M.O.F.-R., B.M., M.T.V.-M. and L.T.; Project administration, L.T. and S.F.; Resources, L.T. and S.F.; Software, U.J., F.P.-V. and L.T.; Supervision, S.F.; Validation, P.R.; Writing—original draft, P.R. and

L.T.; Writing—review & editing, B.G.-d.T., U.J., F.P.-V., M.O.F.-R., A.R., B.M., M.T.V.-M., J.M.-Z., A.C., L.T. and S.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Consejo Nacional de Ciencia y Tecnología (CONACyT), grant number SALUD-2014-1-233721, and by Instituto Nacional de Pediatría, Recursos Fiscales E022 project INP 2014/41.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Instituto Nacional de Pediatría (INP), Ciudad de Mexico, Mexico (protocol code: INP 2014/041; date of approval: 17 July 2014).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study to publish this paper.

Data Availability Statement: Available upon request, following the approved ethics committee guidelines.

Acknowledgments: We would like to thank the patients for their consent to participate in this study. We would like to acknowledge Yañez J., Paul Gaytan P., Lopez E., and Becerra S. from the Sequencing and Synthesis Unit of the Instituto de Biotecnología, UNAM for primer synthesis and sequencing services. Pedro Reyes acknowledges the support of CONACYT, through the scholarship 899784, to develop his PhD in Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

1KGP	1000 Genomes Project
ACMG	American College of Medical Genetics
AMR	Ad Mixed American (1KGP super population)
ANC	Absolut Neutrophils Counts
bp	base pair
BMF	Bone marrow failure
CNS	Central nervous system
CEU	Utah residents (CEPH) with Northern and Western European ancestry (1KGP population)
ChAS	Chromosome Analysis Suite
CHB	Han Chinese in Beijing, China (1KGP population)
DEB	Diepoxibutane
FA	Fanconi anemia
FA/BRCA	Fanconi anemia/Breast Cancer pathway
FPV	Founder pathogenic variant
Hb	Hemoglobin
HGVS	Human Genome Variation Society
HSCT	Hematopoietic stem-cell transplantation
HSF	Human Splicing Finder
Hg19	Homo sapiens (human) genome assembly GRCh37
INP	Instituto Nacional de Pediatría
kb	kilobases
LCSH	Long-contiguous stretches of homozygosity
MCV	Mean Corpuscular Volume
MRI	Magnetic resonance imaging
mRNA	messenger RNA
MXL	Mexican Ancestry in Los Angeles CA, USA (1 KGP population)
MT	Mutant
NAT	Native Americans (1KGP population)
NGS	Next generation sequencing
NMD	Non-sense mediated decay
OMIM	Online Mendelian Inheritance in Man

PCR	Polymerase chain reaction
PHENOS	Skin Pigmentation, small Head, small Eyes, Nervous system, Otology, Short stature
PLT	Platelets
PVS1	Pathogenic Very Strong 1
PM2	Pathogenic Moderate 2
PP3	Pathogenic Supporting 3
PDB	Protein Data Bank
PV	Pathogenic variants
SNV	Single nucleotide variant
SDS	Standard deviation score
TPR	Tetratricopeptide repeat
URTI	Upper respiratory tract infection
VACTERL-H	Vertebral, Anal, Cardiac, Tracheo-esophageal fistula, Esophageal atresia, Renal, upper Limb, and Hydrocephalus
WG	Weeks of gestation
WT	Wild type
y.o.	Year old
YRI	Yoruba in Ibadan, Nigeria (1KGP population)

References

- Auerbach, A.D. Fanconi anemia and its diagnosis. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **2009**, *668*, 4–10. [[CrossRef](#)]
- Alter, B.P.; Giri, N.; Savage, S.; Rosenberg, P. Cancer in the National Cancer Institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. *Haematologica* **2017**, *103*, 30–39. [[CrossRef](#)]
- Savage, S.A.; Walsh, M.F. Myelodysplastic Syndrome, Acute Myeloid Leukemia, and Cancer Surveillance in Fanconi Anemia. *Hematol. Clin. N. Am.* **2018**, *32*, 657–668. [[CrossRef](#)]
- Kottemann, M.C.; Smogorzewska, A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* **2013**, *493*, 356–363. [[CrossRef](#)]
- Rodriguez, A.; D’Andrea, A. Fanconi anemia pathway. *Curr. Biol.* **2017**, *27*, R986–R988. [[CrossRef](#)] [[PubMed](#)]
- Sroka, I.; Frohnmayer, L.; Van Ravenhorst, S.; Wirkkula, L. Fanconi Anemia Clinical Care Guidelines, 5th ed. Available online: https://www.fanconi.org/images/uploads/other/Fanconi_Anemia_Clinical_Care_Guidelines_5thEdition_web.pdf (accessed on 24 December 2021).
- Niraj, J.; Färkkilä, A.; D’Andrea, A.D. The Fanconi Anemia Pathway in Cancer. *Annu. Rev. Cancer Biol.* **2019**, *3*, 457–478. [[CrossRef](#)]
- Gille, J.J.P.; Floor, K.; Kerkhoven, L.; Ameziane, N.; Joenje, H.; De Winter, J.P. Diagnosis of Fanconi Anemia: Mutation Analysis by Multiplex Ligation-Dependent Probe Amplification and PCR-Based Sanger Sequencing. *Anemia* **2012**, *2012*, 603253. [[CrossRef](#)] [[PubMed](#)]
- Callén, E.; Casado, J.A.; Tischkowitz, M.D.; Bueren, J.A.; Creus, A.; Marcos, R.; Dasí, A.; Estella, J.M.; Muñoz, A.; Ortega, J.J.; 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*, 1946–1949. [[CrossRef](#)]
- Whitney, M.A.; Saito, H.; Jakobs, P.M.; Gibson, R.; Moses, R.E.; Grompe, M. A common mutation in the FACC gene causes Fanconi anaemia in Ashkenazi Jews. *Nat. Genet.* **1993**, *4*, 202–205. [[CrossRef](#)] [[PubMed](#)]
- De Vries, Y.; Lwiwski, N.; Levitus, M.; Kuyt, B.; Israels, S.; Arwert, F.; Zwaan, M.; Greenberg, C.R.; Alter, B.P.; Joenje, H.; et al. A Dutch Fanconi Anemia FANCC Founder Mutation in Canadian Manitoba Mennonites. *Anemia* **2012**, *2012*, 865170. [[CrossRef](#)]
- Teresa, B.G.; Frias, S.; Molina, T.V.; Villarreal, M.T.; Rodriguez, A.; Carnevale, A.; López-Hernández, G.; Vollbrechtshausen, L.; Olaya-Vargas, A.; Torres, L. FANCC Dutch founder mutation in a Mennonite family from Tamaulipas, México. *Mol. Genet. Genom. Med.* **2019**, *7*, e710. [[CrossRef](#)]
- Auerbach, A.D.; Greenbaum, J.; Pujara, K.; Batish, S.D.; Bitencourt, M.A.; Kokemohr, I.; Schneider, H.; Lobitz, S.; Pasquini, R.; Giampietro, P.F.; et al. Spectrum of sequence variation in the FANCG gene: An International Fanconi Anemia Registry (IFAR) study. *Hum. Mutat.* **2003**, *21*, 158–168. [[CrossRef](#)] [[PubMed](#)]
- Morgan, N.V.; Essop, F.; Demuth, I.; de Ravel, T.; Jansen, S.; Tischkowitz, M.; Lewis, C.M.; Wainwright, L.; Poole, J.; Joenje, H.; et al. A common Fanconi anemia mutation in black populations of sub-Saharan Africa. *Blood* **2005**, *105*, 3542–3544. [[CrossRef](#)]
- Wilson, J.B.; Blom, E.; Cunningham, R.; Xiao, Y.; Kupfer, G.M.; Jones, N.J. Several tetratricopeptide repeat (TPR) motifs of FANCG are required for assembly of the BRCA2/D1-D2-G-X3 complex, FANCD2 monoubiquitylation and phleomycin resistance. *Mutat. Res. Mol. Mech. Mutagen.* **2010**, *689*, 12–20. [[CrossRef](#)] [[PubMed](#)]
- Blom, E.; van de Vrugt, H.J.; de Vries, Y.; de Winter, J.P.; Arwert, F.; Joenje, H. Multiple TPR motifs characterize the Fanconi anemia FANCG protein. *DNA Repair* **2003**, *3*, 77–84. [[CrossRef](#)] [[PubMed](#)]
- Shakeel, S.; Rajendra, E.; Alcón, P.; O’Reilly, F.J.; Chorev, D.; Maslen, S.; Degliesposti, G.; Russo, C.J.; He, S.; Hill, C.H.; et al. Structure of the Fanconi anaemia monoubiquitin ligase complex. *Nature* **2019**, *575*, 234–237. [[CrossRef](#)] [[PubMed](#)]

18. Ceccaldi, R.; Sarangi, P.; D'Andrea, A.D. The Fanconi anaemia pathway: New players and new functions. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 337–349. [[CrossRef](#)]
19. García-De-Teresa, B.; Rodríguez, A.; Frias, S. Chromosome Instability in Fanconi Anemia: From Breaks to Phenotypic Consequences. *Genes* **2020**, *11*, 1528. [[CrossRef](#)]
20. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–423. [[CrossRef](#)]
21. Li, L.-H.; Ho, S.-F.; Chen, C.-H.; Wei, C.-Y.; Wong, W.-C.; Li, L.-Y.; Hung, S.-I.; Chung, W.-H.; Pan, W.-H.; Lee, M.T.M.; et al. Long contiguous stretches of homozygosity in the human genome. *Hum. Mutat.* **2006**, *27*, 1115–1121. [[CrossRef](#)]
22. Pajusalu, S.; Žilina, O.; Yakoreva, M.; Tammur, P.; Kuuse, K.; Mölter-Väär, T.; Nõukas, M.; Reimand, T.; Öunap, K. The Diagnostic Utility of Single Long Contiguous Stretches of Homozygosity in Patients without Parental Consanguinity. *Mol. Syndr.* **2015**, *6*, 135–140. [[CrossRef](#)]
23. Wang, J.-C.; Ross, L.; Mahon, L.W.; Owen, R.; Hemmat, M.; Wang, B.T.; El Naggar, M.; A Kopita, K.; Randolph, L.M.; Chase, J.M.; et al. Regions of homozygosity identified by oligonucleotide SNP arrays: Evaluating the incidence and clinical utility. *Eur. J. Hum. Genet.* **2014**, *23*, 663–671. [[CrossRef](#)] [[PubMed](#)]
24. Chaves, T.F.; Oliveira, L.F.; Ocampos, M.; Barbato, I.T.; De Luca, G.R.; Filho, J.H.B.; Pinto, L.L.D.C.; Bernardi, P.; Maris, A.F. Long contiguous stretches of homozygosity detected by chromosomal microarrays (CMA) in patients with neurodevelopmental disorders in the South of Brazil. *BMC Med. Genom.* **2019**, *12*, 50. [[CrossRef](#)]
25. Silva-Zolezzi, I.; Hidalgo-Miranda, A.; Estrada-Gil, J.; Fernandez-Lopez, J.C.; Uribe-Figueroa, L.; Contreras, A.; Balam-Ortiz, E.; del Bosque-Plata, L.; Velazquez-Fernandez, D.; Lara, C.; et al. Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8611–8616. [[CrossRef](#)] [[PubMed](#)]
26. Spear, M.L.; Diaz-Papkovich, A.; Ziv, E.; Yracheta, J.M.; Gravel, S.; Torgerson, D.G.; Hernandez, R.D. Recent shifts in the genomic ancestry of Mexican Americans may alter the genetic architecture of biomedical traits. *eLife* **2020**, *9*, e56029. [[CrossRef](#)]
27. Yagasaki, H.; Hamanoue, S.; Oda, T.; Nakahata, T.; Asano, S.; Yamashita, T. Identification and characterization of novel mutations of the major Fanconi anemia gene FANCA in the Japanese population. *Hum. Mutat.* **2004**, *24*, 481–490. [[CrossRef](#)]
28. Demuth, I.; Wlodarski, M.; Tipping, A.J.; Morgan, N.; De Winter, J.P.; Thiel, M.; Gräsl, S.; Schindler, D.; D'Andrea, A.D.; Altay, C.; et al. Spectrum of mutations in the Fanconi anaemia group G gene, FANCG/XRCC9. *Eur. J. Hum. Genet.* **2000**, *8*, 861–868. [[CrossRef](#)]
29. Park, J.; Kim, M.; Jang, W.; Chae, H.; Kim, Y.; Chung, N.-G.; Lee, J.-W.; Cho, B.; Jeong, D.-C.; Park, I.Y.; et al. Founder Haplotype Analysis of Fanconi Anemia in the Korean Population Finds Common Ancestral Haplotypes for a FANCG Variant: Ancestral Haplotypes for AFANCGVariant. *Ann. Hum. Genet.* **2015**, *79*, 153–161. [[CrossRef](#)]
30. Alter, B.P.; Giri, N. Thinking of VACTERL-H? Rule out Fanconi Anemia According to PHENOS: VACTERL-H, PHENOS, and Fanconi Anemia. *Am. J. Med. Genet. Part A* **2016**, *170*, 1520–1524. [[CrossRef](#)]
31. Aguilar-Ordoñez, I.; Pérez-Villatoro, F.; García-Ortiz, H.; Barajas-Olmos, F.; Ballesteros-Villascán, J.; González-Buenfil, R.; Fresno, C.; Garcíarrubio, A.; Fernández-López, J.C.; Tovar, H.; et al. Whole genome variation in 27 Mexican indigenous populations, demographic and biomedical insights. *PLoS ONE* **2021**, *16*, e0249773. [[CrossRef](#)] [[PubMed](#)]
32. Torres-Cisneros, G. Mixes. In *Pueblos Indígenas del Mexico Contemporaneo*; Comisión Nacional para el Desarrollo de los Pueblos Indígenas, PNUD: Mexico City, Mexico, 2004; pp. 1–43.
33. Quinto-Cortés, C.D.; Arriola, L.A.; García-Hughes, G.; García-López, R.; Molina, D.P.; Flores, M.; Palacios, R.; Piñero, D. Genetic Characterization of Indigenous Peoples from Oaxaca, Mexico, and Its Relation to Linguistic and Geographic Isolation. *Hum. Biol.* **2010**, *82*, 409–432. [[CrossRef](#)]
34. Cuentame INEGI. Available online: http://www.cuentame.inegi.org.mx/monografias/informacion/oax/territorio/div_municipal.aspx?tema=me&e=20 (accessed on 8 December 2021).
35. Orphanet: Anemia de Fanconi. Available online: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Lng=ES&Expert=84 (accessed on 19 December 2021).
36. Rogers, K.J.; Fu, W.; Akey, J.M.; Monnat, J.R., Jr. Global and disease-associated genetic variation in the human Fanconi anemia gene family. *Hum. Mol. Genet.* **2014**, *23*, 6815–6825. [[CrossRef](#)]
37. Shimamura, A.; Alter, B.P. Pathophysiology and management of inherited bone marrow failure syndromes. *Blood Rev.* **2010**, *24*, 101–122. [[CrossRef](#)] [[PubMed](#)]
38. Fiesco-Roa, M.O.; Giri, N.; McReynolds, L.J.; Best, A.F.; Alter, B.P. Genotype-phenotype associations in Fanconi anemia: A literature review. *Blood Rev.* **2019**, *37*, 100589. [[CrossRef](#)] [[PubMed](#)]
39. Graf, C.M.; Nichele, S.; Siviero, R.B.; Loth, G.; Trennepohl, J.P.; Zinher, M.T.; Grandinetti, A.; Pilonetto, D.V.; Pasquini, R.; Moreira, A.T.R.; et al. Ocular Manifestations in Patients with Fanconi Anemia: A Single Center Experience Including 106 Patients. *J. Pediatr.* **2021**. [[CrossRef](#)] [[PubMed](#)]
40. Jung, M.; Mehta, P.A.; Jiang, C.S.; Rosti, R.O.; Usleaman, G.; da Rosa, J.M.C.; Lach, F.P.; Goodridge, E.; Auerbach, A.D.; Davies, S.M.; et al. Comparison of the clinical phenotype and haematological course of siblings with Fanconi anaemia. *Br. J. Haematol.* **2020**, *193*, 971–975. [[CrossRef](#)] [[PubMed](#)]
41. Faivre, L.; Guardiola, P.; Lewis, C.; Dokal, I.; Ebell, W.; Zatterale, A.; Altay, C.; Poole, J.; Stones, D.; Kwee, M.L.; et al. Association of complementation group and mutation type with clinical outcome in fanconi anemia. European Fanconi Anemia Research Group. *Blood* **2000**, *96*, 4064–4070.

42. Feben, C.; Kromberg, J.; Wainwright, R.; Stones, D.; Sutton, C.; Poole, J.; Haw, T.; Krause, A. Phenotypic consequences in black South African Fanconi anemia patients homozygous for a founder mutation. *Genet. Med.* **2014**, *16*, 400–406. [[CrossRef](#)]
43. Dillon, B.; Feben, C.; Segal, D.; Du Plessis, J.; Reynders, D.; Wainwright, R.; Poole, J.; Krause, A. Endocrine profiling in patients with Fanconi anemia, homozygous for a FANCG founder mutation. *Mol. Genet. Genom. Med.* **2020**, *8*, e1351. [[CrossRef](#)]
44. Köhler, S.; Gargano, M.; Matentzoglou, N.; Carmody, L.C.; Lewis-Smith, D.; Vasilevsky, N.A.; Danis, D.; Balagura, G.; Baynam, G.; Brower, A.M.; et al. The Human Phenotype Ontology in 2021. *Nucleic Acids Res.* **2021**, *49*, D1207–D1217. [[CrossRef](#)]
45. Esmer, C.; Sánchez, S.; Ramos, S.; Molina, B.; Frias, S.; Carnevale, A. DEB test for Fanconi anemia detection in patients with atypical phenotypes. *Am. J. Med. Genet. Part A* **2005**, *124A*, 35–39. [[CrossRef](#)]
46. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; Del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **2011**, *43*, 491–498. [[CrossRef](#)]
47. Van Der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; Del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinform.* **2013**, *43*, 11.10.1–11.10.33. [[CrossRef](#)]
48. den Dunnen, J.T.; Dalgleish, R.; Maglott, D.R.; Hart, R.K.; Greenblatt, M.S.; McGowan-Jordan, J.; Roux, A.-F.; Smith, T.; Antonarakis, S.E.; Taschner, P.E.; et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum. Mutat.* **2016**, *37*, 564–569. [[CrossRef](#)]
49. Kopanos, C.; Tsiolkas, V.; Kouris, A.; Chapple, C.E.; Aguilera, M.A.; Meyer, R.; Massouras, A. VarSome: The human genomic variant search engine. *Bioinformatics* **2019**, *35*, 1978–1980. [[CrossRef](#)] [[PubMed](#)]
50. Green, M.R.; Sambrook, J. The Basic Polymerase Chain Reaction (PCR). *Cold Spring Harb. Protoc.* **2018**, *5*, pdb-prot095117. [[CrossRef](#)] [[PubMed](#)]
51. Kearney, H.M.; Kearney, J.B.; Conlin, L.K. Diagnostic Implications of Excessive Homozygosity Detected by SNP-Based Microarrays: Consanguinity, Uniparental Disomy, and Recessive Single-Gene Mutations. *Clin. Lab. Med.* **2011**, *31*, 595–613. [[CrossRef](#)]



<https://doi.org/10.18233/APM44No1pp29-552548>

Anemia de Fanconi, Parte 2. Estrategia metodológica para el diagnóstico molecular en pacientes con anemia de Fanconi

Fanconi anemia, Part 2. Methodological strategy for molecular diagnosis in patients with Fanconi anemia

Leda Torres,¹ Ulises Juárez,^{1,2} Pedro Reyes,^{1,2} Sara Frías^{1,3}

Resumen

La anemia de Fanconi (AF) es una enfermedad rara, se presenta en 1-5/millón de nacidos vivos. A nivel celular presentan inestabilidad cromosómica, que es la base para su diagnóstico y aunque clínicamente son heterogéneos, hay tres características generales: alteraciones del desarrollo físico, pancitopenia y alto riesgo a desarrollar cáncer. Presenta heterogeneidad genética, hasta ahora se han reportado 22 genes responsables de la AF, 20 de estos genes se heredan de manera autosómica recesiva, uno autosómica dominante y uno ligada al X, sin embargo, existen genes por detectar, ya que a pesar de una minuciosa búsqueda, no en todos los pacientes se logra encontrar la variante patogénica responsable. Debido a esta heterogeneidad, el diagnóstico molecular es complicado, por lo que se necesita una estrategia con varias metodologías como el ensayo de amplificación de sondas dependiente de ligandos múltiples (MLPA), secuenciación de nueva generación, ya sea por panel dirigido (16 genes *FANC*), o por secuenciación del exoma completo, y microarreglos de alta resolución. Con estas metodologías se pueden detectar grandes deleciones o duplicaciones en los genes *FANC*, alteraciones puntuales y en el número de copias, así como regiones largas con homocigosidad, con el propósito de encontrar alelos homocigotos. En este artículo presentamos una estrategia detallada para realizar la genotipificación de los pacientes AF mexicanos, con un porcentaje de éxito del 80%.

PALABRAS CLAVE: Anemia de Fanconi; Variantes patogénicas; genes *FANC*; MLPA; Secuenciación de nueva generación; Estrategia de genotipificación.

Abstract

Fanconi anemia (FA) is a rare disease occurring in 1-5/million live births. Patients present chromosomal instability at the cellular level, which is the basis for their diagnosis, and although clinically they are heterogeneous, there are three general characteristics: alterations in physical development, pancytopenia and high risk of cancer development. To date, 22 genes responsible for FA have been reported, 20 of which are inherited in an autosomal recessive pattern, one autosomal dominant and one X-linked; however, there are genes to be detected, since despite a thorough search, the responsible pathogenic variant cannot be found in all patients. Due to this heterogeneity, the molecular diagnosis is complicated, so a strategy with several methodologies, such as multiple ligand-dependent probe amplification assay (MLPA) and next-generation sequencing, either by directed panel (16 *FANC* genes) or by whole exome sequencing and high-resolution microarrays, is necessary. With these methodologies, it is possible to detect long deletions or duplications in *FANC* genes, single nucleotide and copy number alterations, and long regions with homozygosity to find homozygous alleles. In this article, we present a detailed strategy for genotyping Mexican FA patients, with a success rate of 80%.

KEYWORDS: Fanconi anemia; Pathogenic variants; *FANC* genes; MLPA; Next generation sequencing; Genotyping strategy.

¹ Laboratorio de Citogenética, Instituto Nacional de Pediatría.

² Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México.

³ Departamento de Medicina Genómica y Toxicología Ambiental, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México.

ORCID

0000-0001-8355-5818
0000-0001-6333-5738
0000-0002-3097-6368

Recibido: 23 de junio de 2022

Aceptado: 02 de enero de 2023

Correspondencia

Sara Frías
sarafrias@iibiomedicas.unam.mx

Este artículo debe citarse como: Torres L, Juárez U, Reyes P, Frías S. Anemia de Fanconi, Parte 2. Estrategia metodológica para el diagnóstico molecular en pacientes con anemia de Fanconi. Acta Pediatr Mex 2023; 44 (1): 29-55.

INTRODUCCIÓN

La anemia de Fanconi (AF) es una enfermedad hereditaria rara, caracterizada por inestabilidad genómica; el fenotipo clínico consiste en anomalías del desarrollo físico, insuficiencia de la médula ósea y predisposición a cáncer. La AF tiene heterogeneidad genética, se han identificado 22 genes *FANC* (**Cuadro 1**). Los productos de estos genes cooperan en la vía FA/BRCA, que regula la respuesta a los agentes genotóxicos y mantiene la integridad del genoma.¹ Todas las variantes patogénicas (VP) reportadas en

estos genes tienen herencia autosómica recesiva, excepto en *FANCB*, que están ligadas al cromosoma X, y en *FANCR/RAD51*, que son autosómicas dominantes.²

Se ha identificado una amplia variedad de mutaciones diferentes entre los genes de la vía FA/BRCA; las VP bialélicas en *FANCA* son la causa más frecuente de AF (60% a 70%), seguidas de las encontradas en *FANCC* (10-12%) y *FANCG* (6-8%)³. El porcentaje de individuos con VP en cada gen *FANC* se muestra en la **Figura 1** (Fanconi Anemia Mutation Database [**Cuadro 1.** Genes implicados en la anemia de Fanconi y localización cromosómica](https://www2.</p>
</div>
<div data-bbox=)

Gene/Alias	Localización Citogenética
<i>FANCA</i>	16q24.3
<i>FANCB</i>	Xp22.2
<i>FANCC</i>	9q22.32
<i>FANCD1/BRCA2</i>	13q13.1
<i>FANCD2</i>	3p25.3
<i>FANCE</i>	6p21.31
<i>FANCF</i>	11p14.3
<i>FANCG/XRCC9</i>	9p13.3
<i>FANCI</i>	15q26.1
<i>FANCI/BRIP1</i>	17q23.2
<i>FANCL</i>	2p16.1
* <i>FANCM</i>	14q21.2
<i>FANCN/PALB2</i>	16q12.2
* <i>FANCO/RAD51C</i>	17q22
<i>FANCP/SLX4</i>	16p13.3
<i>FANCQ/XPF</i>	16p13.12
* <i>FANCR/RAD51</i>	15q15.1
* <i>FANCS/BRCA1</i>	17q21.31
<i>FANCT/UBE2T</i>	1q32.1
<i>FANCU/XRCC2</i>	7q36.1
<i>FANCV/REV7</i>	1p36.22
<i>FANCW/RFWD3</i>	16q23.1

* Estos genes se denominan "Fanconi anemia-like" porque los pacientes que portan una VP en cualquiera de ellos no presentan insuficiencia medular¹

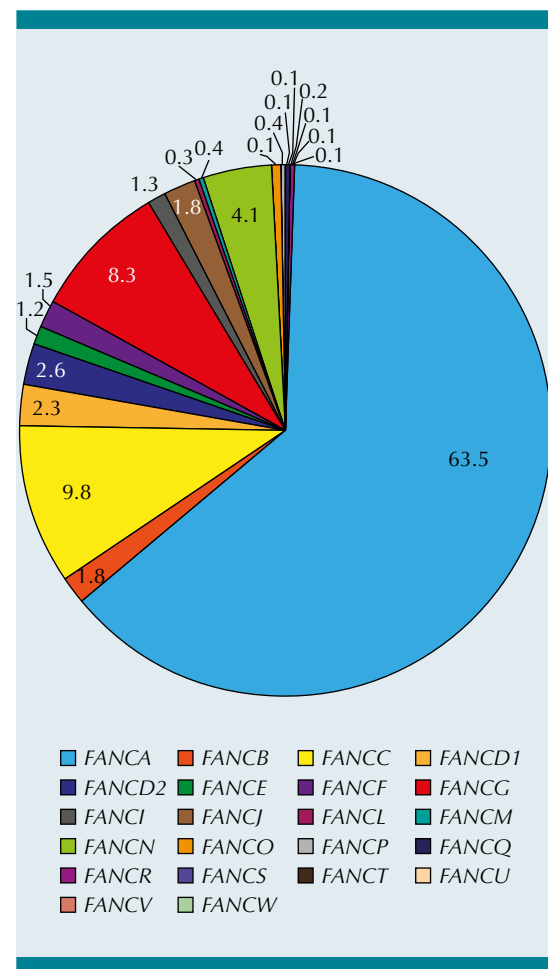


Figura 1. Distribución de pacientes con anemia de Fanconi según el gen *FANC* afectado.

rockefeller.edu/fanconi/; consultada el 20 de enero de 2022).

El hecho de que varios genes y cientos de VP únicas estén asociados con el fenotipo AF, hace que la heterogeneidad de *locus* y de alelo sea la regla. Debido a esta complejidad, el diagnóstico molecular de la enfermedad puede ser desafiante. Distintos grupos de trabajo en el mundo han seguido estrategias diversas de acuerdo a su población y recursos disponibles.⁴⁻⁶ Encontrar la estrategia molecular adecuada para estudiar un grupo de pacientes con AF, puede ser un trabajo largo y costoso.

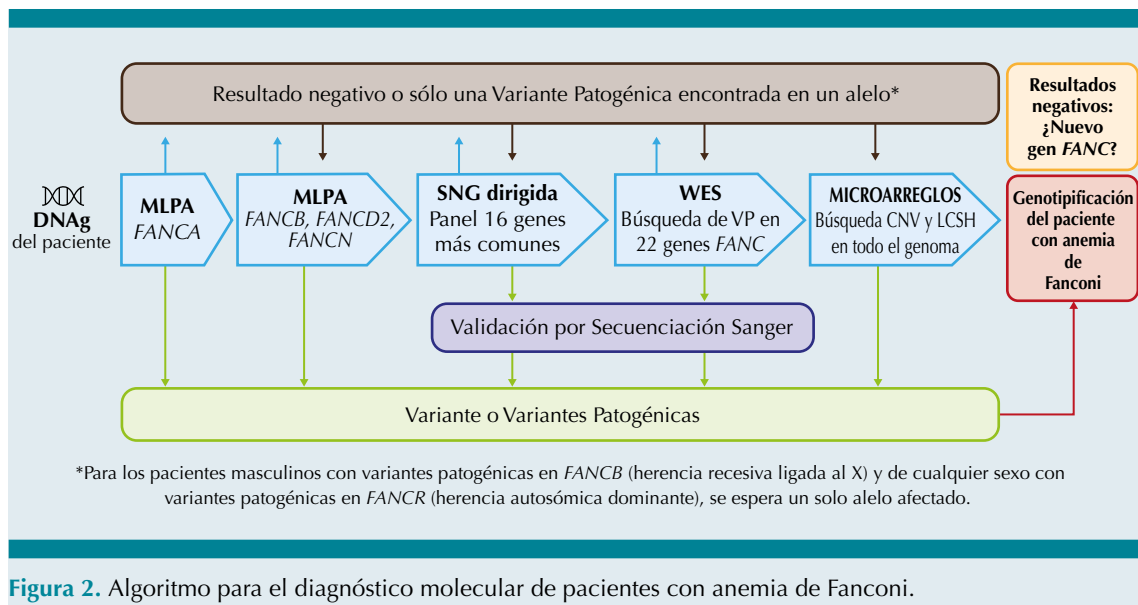
El objetivo de este artículo es presentar la estrategia seguida por nuestro grupo para determinar el genotipo de 50 pacientes con AF, en 40 de los cuales (80%), pudimos obtener el genotipo causal de la enfermedad. Esperamos que este trabajo contribuya a encontrar la estrategia de genotipificación de pacientes con AF más económica y efectiva, en laboratorios con características similares a las nuestras.

En breve, partimos de muestras biológicas de pacientes con diagnóstico clínico y citogenético de AF. Usualmente, las muestras se obtienen de sangre periférica o de aspirado de médula ósea, de las que se extrae DNA genómico (DNAG). A partir de entonces, procedemos de acuerdo con el flujo de trabajo que se sugiere a continuación (**Figura 2**):

1. El primer abordaje es la detección de deleciones largas en el gen *FANCA*, a través de la amplificación de sondas dependiente de ligandos múltiples (MLPA, por sus siglas en inglés), esto debido a que este tipo de VP se observa con mayor frecuencia en este gen, involucrando uno o varios exones, o la pérdida del gen.
2. Para pacientes negativos a deleciones en *FANCA*, realizamos análisis por MLPA para los genes *FANCD2*, *FANCN*

(*PALB2*) y *FANCB*, que también presentan grandes deleciones y duplicaciones detectables por MLPA.

3. Si no se determinan VP mediante MLPA, o si sólo se logró detectar VP en uno de los dos alelos, continuamos la genotipificación a través de secuenciación de nueva generación (NGS, por sus siglas en inglés), ya sea por panel dirigido (en nuestro caso, compuesto por los 16 genes *FANC* más frecuentemente afectados), o por secuenciación del exoma completo (WES, por sus siglas en inglés), para detectar VP pequeñas o puntuales. Con esta estrategia hemos logrado obtener la mayoría de las VP presentes en nuestros pacientes.
4. Si todavía tenemos pacientes sin genotipo, el método de tamizaje que hemos seguido son los microarreglos de alta resolución, que constan de marcadores de polimorfismos de un sólo nucleótido (SNP, por sus siglas en inglés) a lo largo de todo el genoma y que nos permiten detectar variaciones en el número de copias (CNV, por sus siglas en inglés), grandes deleciones o duplicaciones en los genes *FANC*, y la detección de regiones con tramos largos contiguos de homocigosidad (LCSH, por sus siglas en inglés), con el propósito de encontrar alelos homocigotos ininterrumpidos con estado de número de copias neutro de 2 a nivel genómico. Los LCSH se observan en todo el genoma humano, se originan por endogamia (consanguinidad o aislamiento reproductivo, tanto geográfico como social) o por fuerzas evolutivas. Son típicas de las poblaciones consanguíneas, sin embargo, las LCSH también son comunes y en ocasiones, inesperadamente largas en el genoma de personas no consanguíneas;⁷⁻¹⁰ la importancia de su detección es que podrían indicar



la presencia de genes recesivos y si se detectan en alguna zona que contenga alguno de los 20 genes *FANC* con herencia autosómica recesiva, se puede hacer una búsqueda intencionada para detectar VP en ese locus. **Figura 2**

METODOLOGÍA

Pacientes

Se incluyen pacientes con características clínicas de AF y prueba citogenética de aberraciones cromosómicas positiva. Se explica el procedimiento y se solicita el consentimiento informado por escrito antes de obtener las muestras biológicas.

Extracción de DNA genómico

El aspirado de sangre o de médula ósea debe colectarse en tubos que contengan EDTA como anticoagulante. Una de las características clínicas más frecuentes en pacientes con AF es el bajo recuento de glóbulos blancos en la sangre; en consecuencia, se espera obtener poco material

de partida para la extracción de DNAg. Para la extracción se utiliza el kit Gentra Puregene Blood Kit (QIAGEN, cat 158389, Venlo, Limburg, NL) (<https://www.qiagen.com/> Gentra Puregene Handbook 12/2014 33), siguiendo las recomendaciones del fabricante, con mínimas modificaciones.

Este protocolo purifica DNAg de muestras frescas o congeladas, a partir de 3 ml de sangre completa. El protocolo también se puede utilizar para la purificación de DNAg proveniente de células centrifugadas, capa leucocitaria o médula ósea. Si se procesan muestras de capa leucocitaria, se debe escalar directamente el volumen de reactivos utilizados en proporción al volumen de la muestra de sangre original.

Procedimiento

- a. Colocar 3 ml de sangre completa o médula ósea en un tubo cónico de 15 ml, agregar 9 ml de solución de lisis de glóbulos rojos y mezclar invirtiendo 10 veces.

- b.** Incubar 10 min a temperatura ambiente (15–25°C). Invertir al menos una vez durante la incubación.
- c.** Centrifugar durante 5 min a 2000 Fuerza Centrífuga Relativa (FCR) para obtener los glóbulos blancos; si quedan glóbulos rojos, repetir el paso de lisis de glóbulos rojos.
- d.** Desechar cuidadosamente el sobrenadante por pipeteo o decantación, conservando aproximadamente 200 μ l del líquido residual y el concentrado de los glóbulos blancos obtenidos. Agitar vigorosamente el tubo para resuspender el concentrado celular en el líquido residual; el sedimento debe dispersarse por completo por agitación en vórtex.
- e.** Añadir 3 ml de solución de lisis celular y pipetear hacia arriba y hacia abajo para lisar las células, o agitar vigorosamente durante 10 s e incubar a 37 °C hasta que la solución sea homogénea.
- f.** Añadir 15 μ l de solución de RNAasa A (4 mg/ml) y mezclar por inversión 25 veces. Incubar durante 30 min a 37°C. Luego, agregar 15 μ l de proteinasa K (20 mg/ml) e incubar durante 30 min a 55 °C. Inmediatamente poner en hielo para enfriar la muestra.
- g.** Agregar 1 ml de solución de precipitación de proteínas, agitar vigorosamente con vórtex durante 20 s, e incubar en hielo durante 15 min.
- h.** Centrifugar durante 10 min/4°C a 2000 x FCR. Las proteínas precipitadas deben formar un sedimento compacto de color marrón oscuro. Si el sedimento de proteínas no está compacto, repetir la centrifugación.
- i.** Transferir el sobrenadante del paso anterior a un tubo de cónico nuevo de 15 ml. Asegurarse que el sedimento de proteína no se desprenda durante el vertido. Añadir 3 ml de isopropanol y 1 μ l de glucógeno (20 mg/ml).
- j.** Mezclar invirtiendo suavemente 50 veces hasta que el DNAg sea visible como hilos o como un ovillo. Centrifugar durante 10 min a 2000 x FCR. El DNAg puede ser visible como un pequeño precipitado blanco.
- k.** Con la ayuda de una pipeta, retirar el sobrenadante. Agregar 3 ml de etanol al 70 %, e invertir varias veces para lavar el precipitado de DNAg.
- l.** Centrifugar durante 5 min a 2000 x FCR. Remover el sobrenadante, transferir el sedimento con el líquido restante a un tubo cónico de 1.5 ml y lavar nuevamente con 300 μ l de etanol al 70 %. Centrifugar, retirar el sobrenadante y dejar secar el DNAg en un tubo abierto a temperatura ambiente o 65°C.
- m.** Resuspender el sedimento de DNAg en 50 μ l de amortiguador Tris-EDTA (TE) de baja concentración de EDTA (Tris-HCl 10 mM (pH 8.0) + EDTA 0.1 mM), o en agua libre de nucleasas.
- n.** El DNAg obtenido se puede cuantificar en un espectrofotómetro Nanodrop™ (Nanodrop Technologies, Wilmington, DE). Para los ensayos de NGS, podría ser necesario un método de cuantificación más sensible, por ejemplo, a través de cuantificación fluorométrica (Qubit, Thermofisher).
- o.** Evaluar la integridad del DNAg mediante electroforesis en gel de agarosa estándar, al 1 % en amortiguador Tris Acetato-EDTA (TAE) 1X.

Amplificación de sondas dependiente de ligandos múltiples (MLPA)

El MLPA es un ensayo basado en PCR multiplex para detectar CNV. En este método, sondas específicas de una región genómica hibridan con el DNAg y luego se amplifican mediante PCR, utilizando un sólo par de oligonucleótidos.¹¹ La gran ventaja de este método es que permite la amplificación simultánea y semi-cuantitativa de más de 60 sondas en una sola reacción, para la identificación rápida y precisa de deleciones y/o duplicaciones.¹²

Las VP encontradas en *FANCA* representan alrededor del 70% de los casos notificados en AF, de los cuales 15%-20% corresponden a grandes deleciones.^{13,14} La mayoría de ellas se han encontrado principalmente a través de MLPA, a menudo abarcando múltiples exones.¹⁵ En nuestro grupo, estas alteraciones se buscan con el kit de MLPA desarrollado por MRC Holland (<https://www.mrcholland.com/product/P057>). Además de *FANCA*, hay otros genes *FANC* en los cuales se ha reportado frecuentemente CNV, pudiendo ser evaluadas también a través de esta metodología, como *FANCD2* (~ 3% de los pacientes),^{16,17} *FANCN* (<1% de pacientes)¹⁸ y *FANCB* (~ 2% de los pacientes).¹⁹

Según nuestra experiencia, la cantidad óptima de DNAg que se utiliza por reacción debe ser 80 y 90 ng aproximadamente, diluida en amortiguador TE. Por otra parte, cada mezcla de sondas contiene fragmentos que son utilizados como control de calidad (generalmente 9), para evaluar la desnaturalización (fragmentos "D"), la cantidad de DNA y la eficiencia de la ligación (fragmentos "Q"). Es importante incluir muestras de DNA de referencia, a partir de: a) individuos sanos (tanto hombres como mujeres), b) muestra positiva para AF (por ejemplo, la línea celular FA-A-VU817) y c) muestra negativa (buffer TE sin DNAg); estos controles deben incluirse durante cada corrida. Se sugiere utilizar 3 muestras de referencia por cada 21 muestras.

Se sigue el mismo flujo de trabajo sugerido en el protocolo general de MRC Holland (<https://support.mrcholland.com/downloads/files/mlpa-general-protocol-one-tube>), que se resume en los siguientes pasos:

Procedimiento

- a. Desnaturalización del DNA e hibridación de las sondas: Calentar la muestra de DNA de interés (5 µl) durante 5 minutos a 98°C. Luego, agregar 3 µl de la mezcla de sondas a hibridar (1,5 µl de mezcla de sondas SALSA + 1,5 µl de amortiguador MLPA) a la muestra de DNA. Colocar el tubo durante 1 minuto a 95°C e hibridar durante 16 h a 60°C.
- b. Ligación: Una vez transcurrido el tiempo de hibridación, disminuir la temperatura del termociclador a 54°C. Sin retirar los tubos del equipo, agregar 32 µl de la mezcla de ligación e incubar a la misma temperatura durante 15 min. Para inactivar la Ligasa 65, incubar durante 5 minutos a 98 °C.
- c. Amplificación por PCR: Los tubos de reacción se enfrían a temperatura ambiente y se añaden 10 µl de la mezcla de polimerasa. Iniciar la reacción de PCR, utilizando el programa sugerido (95°C por 30 s, 60°C por 30 s y 72°C por 1 minuto, durante 35 ciclos), con una amplificación final a 72°C por 20 minutos. Guardar los tubos protegidos de la luz hasta la electroforesis capilar.
- d. Electroforesis capilar: los productos de PCR se cargan en un instrumento de electroforesis capilar para separarlos según su longitud. La proporción de los componentes de la mezcla de reacción (producto final de la reacción MLPA, marcador de peso y formamida), así como los parámetros de ejecución electroforética, varían según el instrumento disponible.

- e. **Análisis de datos:** La revisión de datos se realiza con el programa Coffalyzer, desarrollado por MRC Holland (<https://www.mrcholland.com/technology/software/coffalyser-net>). Durante este paso, el número relativo de copias se determina contrastando las alturas relativas de los picos de las sondas de referencia y de las regiones blanco en las muestras de prueba contra las de las muestras de referencia. Es fundamental configurar el estudio de las señales crudas obtenidas con la versión de cada mezcla de sondas y el número de lote con el que se realizó la reacción.

Una vez obtenido el electroferograma, se evalúa la altura de los picos de fluorescencia después de la fragmentación. El nivel de fluorescencia de cada pico en una muestra se compara con la señal observada en las muestras de DNA utilizadas como referencia. Por lo tanto, una delección se visualizará como una disminución en la altura del pico en comparación con el fragmento normal, mientras que una duplicación aumenta el tamaño del pico de ese fragmento. Posteriormente, las señales sin procesar se normalizan para obtener un valor que refleje la relación ajustada entre las señales contrastadas. Este valor, conocido como coeficiente de dosis (DQ), permite discriminar el número de copias. El DQ varía de 0 (delecciones homocigotas) a 2,15 (duplicaciones homocigotas). Los valores entre 0,8 y 1 son considerados normales (conjunto diploide) ²⁰.

Secuenciación de nueva generación (NGS)

La NGS es un término genérico utilizado para describir el uso de diversas tecnologías, cuya característica común es la secuenciación masiva y paralela de múltiples fragmentos de DNA.²¹ La combinación de alto rendimiento y secuenciación de una sola molécula de DNA son características distintivas de la NGS, indepen-

dientemente de la plataforma de secuenciación utilizada.²² La evolución y diversificación de los ensayos de NGS ha permitido superar las limitaciones de la secuenciación de Sanger y generar una enorme cantidad de datos de secuencias de nucleótidos, haciéndolos dependientes de un análisis computacional intensivo para el ensamblaje de las secuencias producidas.²³

Los pasos principales de la NGS de DNA implican la fragmentación del DNA, preparación de bibliotecas, secuenciación, análisis bioinformático, anotación de las variantes encontradas y su interpretación.²⁴ Sin embargo, debe considerarse que la elección de la longitud deseada del fragmento ("lecturas" cortas o largas) depende de la pregunta de investigación y su aplicación clínica.

La aplicación de la NGS ha sido especialmente útil en trastornos con alta variabilidad fenotípica, acelerando la identificación y descubrimiento de nuevas variantes causales, lo que ha impulsado su aplicación en diagnóstico molecular.²⁵ La reducción de costos y el potencial de la información obtenida han permitido la integración de datos de NGS a la práctica clínica en medicina.²⁶ Actualmente, existen diferentes enfoques para NGS, dependiendo de si el alcance involucra una gran cantidad de genes (paneles de genes dirigidos), regiones codificantes de proteínas (WES) o secuenciación del genoma completo (WGS, por sus siglas en inglés).

Según la última versión de la Guía de Atención Clínica de la AF, la NGS debe realizarse como parte del proceso de genotipificación, seleccionando un panel de genes *FANC* en una primera etapa, apelando a WES en caso de resultados negativos por prueba de panel.²⁷ Nuestro grupo ha seguido esta estrategia para la mayoría de los casos analizados, considerando los beneficios y limitaciones de cada enfoque.

El uso de paneles dirigidos se ha generalizado en el contexto de la AF, debido a la mejor

caracterización de la vía FA/BRCA, que ha permitido estudiar más genes *FANC* en un solo ensayo, además de ofrecer mayor profundidad de cobertura a un costo relativamente bajo. Sin embargo, la secuenciación de exoma completo, se ha vuelto cada vez más atractiva debido a que con ella, se pueden descubrir VP en regiones genómicas diferentes de los genes *FANC* conocidos y con esto, descubrir nuevos genes *FANC* o candidatos.²⁷ Un ejemplo que ilustra esta idea es el descubrimiento de VP bialélicas en *ERCC4/XP-F*, en un grupo de pacientes con AF sin clasificación genotípica; la pérdida de función de este gen se había asociado previamente a Xeroderma pigmentoso. Este hallazgo condujo a la reclasificación de este gen como *FANCQ*.²⁸ De manera similar, se descubrieron VP en *FANCT* (también conocido como *UBE2T*, en dos pacientes japoneses)²⁹ y en *FANCU*.³⁰

Nuestra búsqueda de VP en la AF a través de NGS se encuentra resumida en el siguiente flujo de trabajo:

Procedimiento

- a. Diseño del panel de genes. El diseño *in silico* de los oligonucleótidos flanqueadores, específicos para 16 genes *FANC* (*FANCA*, *FANCB*, *FANCC*, *FANCD1/BRCA2*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *FANCI/BRIP1*, *FANCL*, *FANCM*, *FANCN/PALB2*, *FANCO/RAD51C*, *FANCP/SLX4* y *FANCQ/ERCC4*) se realiza con la aplicación web SureDesign (Agilent Technologies). Se espera que el diseño final abarque el 99,5% de todos los exones codificantes de estos genes más 50 pb de sus regiones intrónicas 5' y 3' flanqueantes.
- b. La digestión, síntesis de las bibliotecas y la amplificación por PCR de los fragmentos seleccionados se logra mediante el sistema de enriquecimiento

Haloplex (Agilent Technologies Inc., 2013, Waghäusel-Wiesental, Alemania).

- c. Las lecturas de los extremos cortos pareados son obtenidas mediante secuenciación basada en tecnología de síntesis, en una plataforma Illumina Mi-Seq (Illumina, San Diego, CA, EE. UU.).
- d. En caso de que el análisis bioinformático de los datos de panel sea negativo, se procede a realizar WES, con las adaptaciones metodológicas sugeridas para este enfoque.

A diferencia del diseño de panel, en el caso de WES, existen diversas opciones comerciales enfocadas en la captura de regiones exónicas (por ejemplo, SureSelectXT Human All Exon V4 (Agilent Technologies), KAPA HyperExome (Roche), entre otras). Además, el uso de reactivos y formulaciones de kits compatibles con los sistemas antes mencionados se extiende a los pasos de enriquecimiento y construcción de bibliotecas. Finalmente, las plataformas de secuenciación también se adaptan para un rendimiento más escalable (por ejemplo, el sistema NovaSeq 6000 (Illumina)); debido a la amplificación de un mayor número de dianas de DNA.

Es importante señalar que, debido a la complejidad técnica e interpretativa de la NGS, usualmente su implementación está limitada a laboratorios públicos o privados con capacidad para cumplir con las recomendaciones y requerimientos sugeridos en los estándares internacionales.³¹

Debido a estas demandas, estos laboratorios ofrecen con frecuencia servicios de NGS a investigadores y médicos. Nuestro grupo ha utilizado esta modalidad para obtener datos de secuenciación a partir del envío de muestras de DNA de alta calidad. La selección del enfoque, así como de las plataformas, depende del propósito del proyecto y de los recursos

disponibles para el cribado molecular en los pacientes con AF.

Análisis de datos de NGS

Debido a la gran cantidad de datos generados a partir de la NGS, incluso a partir de paneles pequeños de genes, se requiere de análisis computacionales para interpretar efectivamente la información de las secuencias.²⁵ El conjunto de algoritmos bioinformáticos ejecutados en una sucesión de pasos para procesar los datos iniciales generados a partir de la plataforma NGS se conoce como *pipeline* bioinformático.²¹ A pesar de la heterogeneidad de las plataformas y el alcance del estudio, el flujo de trabajo bioinformático consta de cinco etapas principales: a) control de calidad de los datos sin procesar (a partir de un archivo en formato de texto que contiene las lecturas de las secuencias y los puntajes individuales de calidad de bases, denominado FASTQ), b) pre-procesamiento, c) mapeo y post-alineamiento, d) llamado de variantes, e) anotación y priorización.^{23,32,33}

Desarrollamos un *pipeline* validado para agilizar el análisis computacional y maximizar la precisión en el llamado de variantes, según las buenas prácticas recomendadas en GATK (Genome Analysis Toolkit).³⁴⁻³⁶ Los programas utilizados para la manipulación de archivos a menudo se empaquetan en una colección de herramientas basadas en línea de comandos para analizar datos de secuenciación de alto rendimiento (GATK “*toolkit*”), enfocadas principalmente en el descubrimiento de variantes. Es muy importante resaltar que estas herramientas sólo se ejecutan en sistemas operativos Linux o MacOSX.

Procedimiento (las opciones de programas se nombran entre paréntesis)

- a. Realizar el control de calidad de las lecturas de secuencia (FastQC).

- b. Filtrar y recortar las lecturas que contienen bases de mala calidad, así como adaptadores de secuenciación (Trimmomatic, Cutadapt y Trim Galore).
- c. Las lecturas recortadas son alineadas para mapear con una secuencia de referencia (alineación de Burrows-Wheeler (BWA-MEM), usando las referencias que se encuentran en las bases de datos de UCSC o Ensembl).
- d. Transformación de los archivos SAM en archivos BAM; y marcaje de duplicados (Samtools y Picard).
- e. Comprobar la calidad de los archivos BAM (Qualimap).
- f. Seguir las recomendaciones de GATK, para obtener el archivo de llamado de variantes (VCF), en modo gVCF (genotipado conjunto).
- g. Finalmente, se anotan las variantes (VEP, SnpEff o ANNOVAR). Las variantes se describen siguiendo las pautas propuestas por la nomenclatura de la *Human Genome Variation Society* (HGVS).³⁷

Clasificación de variantes

Una vez filtradas las variantes candidatas por paciente, el siguiente paso consiste en clasificarlas según su grado de patogenicidad. El sistema más utilizado para esta tarea proviene de las recomendaciones sugeridas por el Colegio Americano de Genética Médica y la Asociación de Patología Molecular (ACMG-AMP, por sus siglas en inglés).^{38,39} De acuerdo con esta guía, la clasificación de variantes está basada en cinco clases: benigna, probablemente benigna, variante de significado desconocido o incierto (VUS, por sus siglas en inglés), probablemente patógena y patógena.

Los criterios utilizados para esta clasificación se fundamentan en un marco de evidencia

en el que se consideran diferentes parámetros (funcionales, poblacionales, datos de segregación, algoritmos computacionales y predictivos, entre otros) para establecer un resultado final. Varias herramientas web facilitan el ejercicio de clasificación, ya que incorporan información valiosa de bases de datos (como ExAc, ClinVar, gnomAD, entre otras) en una sola plataforma (por ejemplo, Varsome).

Validación de variantes encontradas por NGS

Debido a que las técnicas de NGS no están libres de errores, la ACMG y la Sociedad Europea de Genética Humana (ESHG, por sus siglas en inglés) recomiendan enfáticamente validar las variantes seleccionadas. Existen factores tanto técnicos como bioinformáticos relacionados con la dificultad para confirmar variantes, principalmente en genes con pseudogenes, regiones repetitivas y con alto contenido de GC o variantes *hotspot*. La secuenciación de Sanger es el método más recomendable para la validación de los resultados de la NGS.

En nuestro caso, se diseñan pares de oligonucleótidos flanqueantes específicos para validar la VP encontrada por NGS, utilizando el algoritmo Primer3 (<http://bioinfo.ut.ee/primer3-0.4.0/>), una de las herramientas de diseño de cebadores más utilizadas.⁴⁰ La amplificación por PCR se lleva a cabo a partir de 100 ng de DNA molde. Los productos de PCR se purifican con el kit QIAquick PCR Purification (QIAGEN, Venlo, Limburg, NL) según las instrucciones del fabricante. Los amplicones purificados se secuencian bidireccionalmente utilizando el kit de secuenciación Big Dye Terminator y se resuelven en un secuenciador Applied Biosystems 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, EE. UU).

Los electroferogramas se analizan con el software Chromas V2.6.6 (www.technelysium.com.au) y las secuencias blanco se comparan con las secuencias de referencia correspondientes

desde GenBank (<https://www.ncbi.nlm.nih.gov>). Existen otros programas alternativos de análisis (<https://www.genewiz.com/Public/Resources/Tools-for-Viewing-Sequencing-Data/>) o aplicaciones web como GEARgenomics (<https://www.gear-genomics.com>, apoyado por EMBL-EBI, EMBL-European Bioinformatics Institute).

La secuenciación de Sanger como estrategia de genotipificación

La secuenciación de Sanger podría realizarse para buscar VP en genes específicos, es decir, cuando se estudian miembros de una familia con una VP previamente identificada o en individuos pertenecientes a una comunidad donde se ha detectado una mutación fundadora, particularmente en poblaciones genéticamente aisladas, como *FANCA* c.295C>T en gitanos españoles,⁴¹ *FANCC* c.456+4A>T en judíos Askenazí⁴², *FANCC* c.67delG en la comunidad menonita⁴³ y *FANCG* c.511-3_511-2delCA en una comunidad Mixe mexicana.⁴⁴

Microarreglos de alta resolución

Realizamos el análisis del DNAG con microarreglos CytoScan 750K (Affymetrix, Santa Clara, CA, EE. UU.). Estos microarreglos proporcionan cobertura de todo el genoma, incluidos 550,000 marcadores para detectar CNV y 200,436 sondas SNP.

Los pasos de digestión, ligación, amplificación por PCR, fragmentación, etiquetado, desnaturalización e hibridación del DNA en el microarreglo, se realizan de acuerdo con los protocolos y pautas de control de calidad proporcionados por el proveedor.

Luego, los microarreglos se tiñen y lavan en la estación de fluidos Affymetrix GeneChip Fluidics Station 450 y se escanean en un Affymetrix GeneChip Scanner 3000 7G (Affymetrix, Santa Clara, CA, EE. UU.); los archivos obtenidos, se analizan con el programa Chromosome Analysis



Suite (ChAS) versión 4.1, proporcionado por Affymetrix (Affymetrix, Santa Clara, CA, EE. UU.) (Anotación NetAffx 33 hg19) (<http://www.affymetrix.com>). Para el análisis del número de copias (CN), se utilizan los parámetros de “alta resolución”: ganancias y pérdidas en 100Kb con ≥ 50 marcadores (sondas).

Análisis de segmentos largos contiguos de homocigosidad (LCSH)

Los umbrales mínimos para los LCSH en análisis clínicos generalmente se establecen entre 3 y 10 Mb.⁴⁵

Los LCSH se visualizan en el programa ChAS, versión 4.1, proporcionado por Affymetrix (Affymetrix, Santa Clara, CA, EE. UU.). Para el análisis, se utilizan los archivos de anotación NetAffx 33 hg19 (<http://www.affymetrix.com>). Para LCSH > 3Mb, la configuración de análisis se establece en LOH con recuento de marcadores (sondas) = 50 y tamaño = 3000 Kbp, y para LCSH > 5Mb se establece en recuento de marcadores = 50 y tamaño = 5000 Kbp.

Reporte de resultados

Una vez clasificadas y validadas las variantes, el siguiente paso es redactar un informe o reporte de resultados. Además de la ACMG, otras organizaciones, como ClinGen, (siglas en inglés de *Clinical Genome Resources*), la ESHG y la Asociación para la Ciencia Genética Clínica (ACGS siglas en inglés de *Association for Clinical Genetic Science*) han desarrollado guías para asesorar a los laboratorios en la notificación y descripción de variantes.^{38,39,46-48} Estos documentos son emitidos por un laboratorio hacia el médico tratante y describen los resultados del análisis genético. Deben ser claros, concisos y seguir las recomendaciones sugeridas en las guías existentes.⁴⁸

Recomendaciones

- a. En general, el formato del informe incluye información sobre el médico solicitante, el laboratorio donde se realizó el análisis y el paciente.
- b. Para la descripción e interpretación de los resultados, la nomenclatura de las variantes se redacta según la Organización del Genoma Humano (HUGO, siglas en inglés de *Human Genome Organization*, <https://www.genenames.org/>) y la Sociedad de Variación del Genoma Humano (HGVS, siglas en inglés de *Human Genome Variation Society* <https://varnomen.hgvs.org/>). El informe debe incluir una interpretación clara de los resultados de la prueba, basada en una revisión del contexto clínico del paciente y la información familiar disponible.⁴⁸ El objetivo debe ser generar una declaración simple e inequívoca que explique las implicaciones clínicas del resultado.
- c. Los detalles metodológicos (principio de la tecnología utilizada, sensibilidad, resolución, limitaciones, herramientas bioinformáticas, entre otros) relevantes para la interpretación de los resultados deben incluirse en el informe. Por ejemplo, en el caso del uso de datos de secuenciación masiva para variantes constitucionales, es recomendable consultar las guías más recientes desarrolladas por la ACMG.
- d. Cada laboratorio debe documentar la evidencia utilizada para clasificar las variantes de acuerdo al impacto que pueden tener en la condición estudiada y establecer políticas precisas para definir cuales variantes se excluirán del informe.³⁹

En 2008, se creó una base de datos de mutaciones de la AF como un esfuerzo cooperativo para acelerar el acceso a la información sobre las variantes que se encuentran en los genes *FANC* (<https://www2.rockefeller.edu/fanconi/mutate/>). La adición de una variante (nueva o previamente notificada) debe hacerse como institución y no de forma individual. Compartir datos con la comunidad científica aumenta significativamente la cantidad de evidencia existente para una variante particular en las bases de datos públicas, y contribuye a mejorar los enfoques en la búsqueda de correlaciones genotipo-fenotipo.⁴⁹ Además, se debe fomentar esta práctica porque facilita la reclasificación de variantes, siendo especialmente importante en el caso de VUS, donde el uso de esta información se suma a la evidencia generada a través de algoritmos bioquímicos predictivos y validación de ensayos funcionales.⁵⁰

- e. Resultados negativos en un paciente con estudio de aberraciones cromosómicas positivo.

Los resultados negativos de las pruebas moleculares deben interpretarse cuidadosamente para un individuo con una prueba de aberraciones cromosómicas dentro del rango de FA. Una explicación para un resultado negativo puede ser la presencia de variantes en un gen *FANC* no descubierto o un tipo de variante en un gen que no se puede identificar con las tecnologías empleadas. Otra posibilidad es el mosaicismo somático.^{51,52} En estos casos, se puede considerar el análisis de un tipo de muestra alternativo a la sangre periférica o médula ósea, como fibroblastos, en individuos que presenten un fenotipo de AF y estudios genéticos negativos en sangre periférica.

DISCUSIÓN

En un porcentaje variable de pacientes (en nuestro caso fue el 20%), no es posible llegar a la genotipificación. Las posibles razones son

a) podrían ser variantes que no se detectan por secuenciación de exoma, como aquellas que se encuentran en regiones intrónicas o regulatorias de los genes y que requerirían otra estrategia más costosa, como la secuenciación del genoma completo; b) alternativamente, podría deberse a que las variantes se encuentren en genes aún no descritos como genes *FANC*, por lo que es recomendable realizar un análisis de los genes asociados funcionalmente a la vía FA/BRCA.

La gran heterogeneidad genética de la AF dificulta la implementación de técnicas de detección clásicamente utilizadas en otro tipo de enfermedades de herencia mendeliana. Por ello, consideramos que la divulgación detallada de los aspectos metodológicos inherentes a la genotipificación en la AF, es fundamental para establecer los requerimientos biológicos y tecnológicos mínimos para identificar las variantes responsables de la enfermedad, en especial en zonas geográficas en donde este abordaje no se ha extendido o se encuentra limitado. Esta explicación, que abarca desde la obtención del DNAG hasta la generación del reporte, pone en contexto la importancia de la información obtenida (como la frecuencia, localización y el tipo de variante) para desarrollar una estrategia de genotipificación costo-efectiva y apropiada a la población AF.

CONCLUSIONES

Describir el genotipo de AF en una población es relevante por varias razones: en primer lugar, mejora la caracterización de los pacientes y por lo tanto, contribuye a la identificación de correlaciones genotipo-fenotipo de AF. En segundo lugar, refina la adaptación de las estrategias de diagnóstico genético en grupos específicos de pacientes, incluidas las selecciones adecuadas de donantes de médula ósea de hermanos sin AF y el asesoramiento genético adecuado. Finalmente, la descripción de casos con genotipos raramente descritos aporta nuevos datos a

considerar cuando se enfrentan a variantes con antecedentes de clasificación de patogenicidad ambigua.

Agradecimientos

Esta investigación fue parcialmente financiada por el Consejo Nacional de Ciencia y Tecnología (CONACyT), número de proyecto SALUD-2014-1-233721; Instituto Nacional de Pediatría, Proyecto Recursos Fiscales E022 INP 2014/41; Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), Universidad Nacional Autónoma de México (UNAM) Grant Number IN205120. Los autores agradecen al M. en C. Francisco Zúñiga-Frias por la revisión del inglés de este manuscrito.

REFERENCIAS

- Fiesco-Roa MÓ, García-de Teresa B, Leal-Anaya P, van 't Hek R, Wegman-Ostrosky T, Frías S, et al. Fanconi anemia and dyskeratosis congenita/telomere biology disorders: Two inherited bone marrow failure syndromes with genomic instability. *Front Oncol* [Internet]. 2022;12. Available from: <https://www.frontiersin.org/articles/10.3389/fonc.2022.949435>
- García-de-Teresa B, Rodríguez A, Frias S. Chromosome Instability in Fanconi Anemia: From Breaks to Phenotypic Consequences. *Genes (Basel)*. 2020 Dec;11(12).
- Niraj J, Färkkilä A, D'Andrea AD. The Fanconi Anemia Pathway in Cancer. *Annu Rev cancer Biol*. 2019 Mar;3:457-78.
- Pilonetto D V, Pereira NF, Bonfim CMS, Ribeiro LL, Bitencourt MA, Kerkhoven L, et al. A strategy for molecular diagnostics of Fanconi anemia in Brazilian patients. *Mol Genet genomic Med*. 2017 Jul;5(4):360-72.
- Mori M, Hira A, Yoshida K, Muramatsu H, Okuno Y, Shiraishi Y, et al. Pathogenic mutations identified by a multimodality approach in 117 Japanese Fanconi anemia patients. *Haematologica*. 2019 Oct;104(10):1962-73.
- Ben Haj Ali A, Messaoud O, Elouej S, Talmoudi F, Ayed W, Mellouli F, et al. FANCA Gene Mutations in North African Fanconi Anemia Patients. *Front Genet* [Internet]. 2021;12. Available from: <https://www.frontiersin.org/articles/10.3389/fgene.2021.610050>
- Li LH, Ho SF, Chen CH, Wei CY, Wong WC, Li LY, et al. Long contiguous stretches of homozygosity in the human genome. *Hum Mutat*. 2006 Nov;27(11):1115-21.
- Pajusalu S, Žilina O, Yakoreva M, Tammur P, Kuuse K, Mõlter-Väär T, et al. The Diagnostic Utility of Single Long Contiguous Stretches of Homozygosity in Patients without Parental Consanguinity. *Mol Syndromol*. 2015 Sep;6(3):135-40.
- Wang JC, Ross L, Mahon LW, Owen R, Hemmat M, Wang BT, et al. Regions of homozygosity identified by oligonucleotide SNP arrays: evaluating the incidence and clinical utility. *Eur J Hum Genet*. 2015 May;23(5):663-71.
- Chaves TF, Oliveira LF, Ocampos M, Barbato IT, de Luca GR, Barbato Filho JH, et al. Long contiguous stretches of homozygosity detected by chromosomal microarrays (CMA) in patients with neurodevelopmental disorders in the South of Brazil. *BMC Med Genomics*. 2019 Mar;12(1):50.
- Stuppia L, Antonucci I, Palka G, Gatta V. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. *Int J Mol Sci*. 2012;13(3):3245-76.
- Hömig-Hölzel C, Savola S. Multiplex ligation-dependent probe amplification (MLPA) in tumor diagnostics and prognostics. *Diagnostic Mol Pathol Am J Surg Pathol part B*. 2012 Dec;21(4):189-206.
- Ameziane N, Errami A, Léveillé F, Fontaine C, de Vries Y, van Spaendonck RML, et al. Genetic subtyping of Fanconi anemia by comprehensive mutation screening. *Hum Mutat*. 2008 Jan;29(1):159-66.
- Castella M, Pujol R, Callén E, Trujillo JP, Casado JA, Gille H, et al. Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. *Blood*. 2011 Apr;117(14):3759-69.
- Kimble DC, Lach FP, Gregg SQ, Donovan FX, Flynn EK, Kamat A, et al. A comprehensive approach to identification of pathogenic FANCA variants in Fanconi anemia patients and their families. *Hum Mutat*. 2018 Feb;39(2):237-54.
- Kalb R, Neveling K, Hoehn H, Schneider H, Linka Y, Batish SD, 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 May;80(5):895-910.
- Flynn EK, Kamat A, Lach FP, Donovan FX, Kimble DC, Narisu N, et al. Comprehensive analysis of pathogenic deletion variants in Fanconi anemia genes. *Hum Mutat*. 2014 Nov;35(11):1342-53.
- Xia B, Dorsman JC, Ameziane N, de Vries Y, Roomans MA, Sheng Q, et al. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat Genet*. 2007 Feb;39(2):159-61.
- Jung M, Ramanagoudr-Bhojappa R, van Twest S, Rosti RO, Murphy V, Tan W, et al. Association of clinical severity with FANCB variant type in Fanconi anemia. *Blood*. 2020 Apr;135(18):1588-602.
- Coffa J, van den Berg J. Analysis of MLPA Data Using Novel Software Coffalyser.NET by MRC-Holland. In: Eldin AB, editor. *Modern Approaches To Quality Control* [Internet]. Rijeka: IntechOpen; 2011. Available from: <https://doi.org/10.5772/21898>
- Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, et al. Standards and Guidelines for Validating

- Next-Generation Sequencing Bioinformatics Pipelines: A Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J Mol Diagn.* 2018 Jan;20(1):4–27.
22. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. *Hum Immunol.* 2021 Nov;82(11):801–11.
 23. Voelkerding K V, Coonrod EM, Durtschi JD, Margraf RL. Next-Generation Sequencing: Principles for Clinical Application BT - Molecular Pathology in Clinical Practice. In: Leonard DGB, editor. Cham: Springer International Publishing; 2016. p. 889–909. Available from: https://doi.org/10.1007/978-3-319-19674-9_59
 24. Qin D. Next-generation sequencing and its clinical application. *Cancer Biol Med.* 2019 Feb;16(1):4–10.
 25. den Dunnen JT. Efficient variant data preparation for Human Mutation manuscripts: Variants and phenotypes. Vol. 40, Human mutation. United States; 2019. p. 1009.
 26. Rizzo JM, Buck MJ. Key principles and clinical applications of “next-generation” DNA sequencing. *Cancer Prev Res (Phila).* 2012 Jul;5(7):887–900.
 27. Sroka I, Frohnmayer L, Wirkkula L. Fanconi Anemia Clinical Care Guidelines [Internet]. Fifth Edit. Vol. 1, Fanconi Anemia Research Fund, Inc. 2020. 1–276 p. Available from: <http://marefateadyan.nashriyat.ir/node/150>
 28. Bogliolo M, Schuster B, Stoecker C, Derkunt B, Su Y, Raams A, et al. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet.* 2013 May;92(5):800–6.
 29. Hira A, Yoshida K, Sato K, Okuno Y, Shiraishi Y, Chiba K, et al. Mutations in the gene encoding the E2 conjugating enzyme UBE2T cause Fanconi anemia. *Am J Hum Genet.* 2015 Jun;96(6):1001–7.
 30. Shamseldin HE, Elfaki M, Alkuraya FS. Exome sequencing reveals a novel Fanconi group defined by XRCC2 mutation. *J Med Genet.* 2012 Mar;49(3):184–6.
 31. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med Off J Am Coll Med Genet.* 2013 Sep;15(9):733–47.
 32. McCombie WR, McPherson JD, Mardis ER. Next-Generation Sequencing Technologies. *Cold Spring Harb Perspect Med.* 2019 Nov;9(11).
 33. Bao R, Huang L, Andrade J, Tan W, Kibbe WA, Jiang H, et al. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Inform.* 2014;13(Suppl 2):67–82.
 34. DePristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011 May;43(5):491–8.
 35. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma.* 2013;43(1110):11.10.1-11.10.33.
 36. Franke KR, Crowgey EL. Accelerating next generation sequencing data analysis: an evaluation of optimized best practices for Genome Analysis Toolkit algorithms. *Genomics Inform.* 2020 Mar;18(1):e10.
 37. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum Mutat.* 2016 Jun;37(6):564–9.
 38. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med Off J Am Coll Med Genet.* 2015 May;17(5):405–24.
 39. Rehder C, Bean LJH, Bick D, Chao E, Chung W, Das S, et al. Next-generation sequencing for constitutional variants in the clinical laboratory, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med Off J Am Coll Med Genet.* 2021 Aug;23(8):1399–415.
 40. Kumar A, Chordia N. In Silico PCR Primer Designing and Validation BT - PCR Primer Design. In: Basu C, editor. New York, NY: Springer New York; 2015. p. 143–51. Available from: https://doi.org/10.1007/978-1-4939-2365-6_10
 41. Callén E, Casado JA, Tischkowitz MD, Bueren JA, Creus A, Marcos R, et al. A common founder mutation in FANCA underlies the world’s highest prevalence of Fanconi anemia in Gypsy families from Spain. *Blood.* 2005 Mar;105(5):1946–9.
 42. Whitney MA, Jakobs P, Kaback M, Moses RE, Grompe M. The Ashkenazi Jewish Fanconi anemia mutation: incidence among patients and carrier frequency in the at-risk population. *Hum Mutat.* 1994;3(4):339–41.
 43. García-de Teresa B, Frias S, Molina B, Villarreal MT, Rodríguez A, Carnevale A, et al. FANCC Dutch founder mutation in a Mennonite family from Tamaulipas, México. *Mol Genet genomic Med.* 2019 Jun;7(6):e710.
 44. Reyes P, García-de Teresa B, Juárez U, Pérez-Villatoro F, Fiesco-Roa MO, Rodríguez A, et al. Fanconi Anemia Patients from an Indigenous Community in Mexico Carry a New Founder Pathogenic Variant in FANCG. *Int J Mol Sci.* 2022 Feb;23(4).
 45. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med Off J Am Coll Med Genet.* 2011 Jul;13(7):680–5.
 46. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, et al. ClinGen—the Clinical Genome Resource. *N Engl J Med.* 2015 Jun;372(23):2235–42.
 47. Matthijs G, Souche E, Alders M, Corveleyn A, Eck S, Feenstra I, et al. Guidelines for diagnostic next-generation sequencing. *Eur J Hum Genet.* 2016 Oct;24(10):1515.

48. Smith K, Martindale J, Wallis Y, Bown N, Leo N, Creswell L, et al. General genetic laboratory reporting recommendations. *Birmingham Assoc Clin Genet Sci*. 2015;11.
49. Directors ABO. Laboratory and clinical genomic data sharing is crucial to improving genetic health care: a position statement of the American College of Medical Genetics and Genomics. *Genet Med Off J Am Coll Med Genet*. 2017 Jul;19(7):721–2.
50. Federici G, Soddu S. Variants of uncertain significance in the era of high-throughput genome sequencing: a lesson from breast and ovary cancers. *J Exp Clin Cancer Res*. 2020 Mar;39(1):46.
51. Gregory JJ, Wagner JE, Verlander PC, Levran O, Batish SD, Eide CR, et al. Somatic mosaicism in Fanconi anemia: evidence of genotypic reversion in lymphohematopoietic stem cells. *Proc Natl Acad Sci U S A*. 2001 Feb;98(5):2532–7.
52. Ikeda H, Matsushita M, Waisfisz Q, Kinoshita A, Oostra AB, Nieuwint AWM, et al. Genetic reversion in an acute myelogenous leukemia cell line from a Fanconi anemia patient with biallelic mutations in BRCA2. *Cancer Res*. 2003 May;63(10):2688–94.

INTRODUCTION

Fanconi anemia (FA) is a rare hereditary disease characterized by genomic instability. The clinical phenotype consists of developmental abnormalities, bone marrow failure, and cancer predisposition. FA presents genetic heterogeneity, 22 *FANCA* genes have been identified to date (**Table 1**). The products of these genes cooperate in the FA/BRCA pathway, which regulates the response to genotoxic agents and maintains genome integrity.¹ All pathogenic variants (PV) in these genes have an autosomal recessive inheritance, except in *FANCB*, which is linked to the X chromosome, and *FANCR/RAD51*, which is autosomal dominant.²

Many different mutations have been identified among the FA/BRCA pathway genes. These include the biallelic PV in *FANCA* as the most frequent cause of FA (60% to 70%), followed by *FANCC* (10-12%) and *FANCG* (6-8%).³ The percentage of individuals with PV in each FA gene is shown in **Figure 1** (Fanconi Anemia Mutation Database <https://www2.rockefeller.edu/fanconi/>; consulted 20th January 2022).

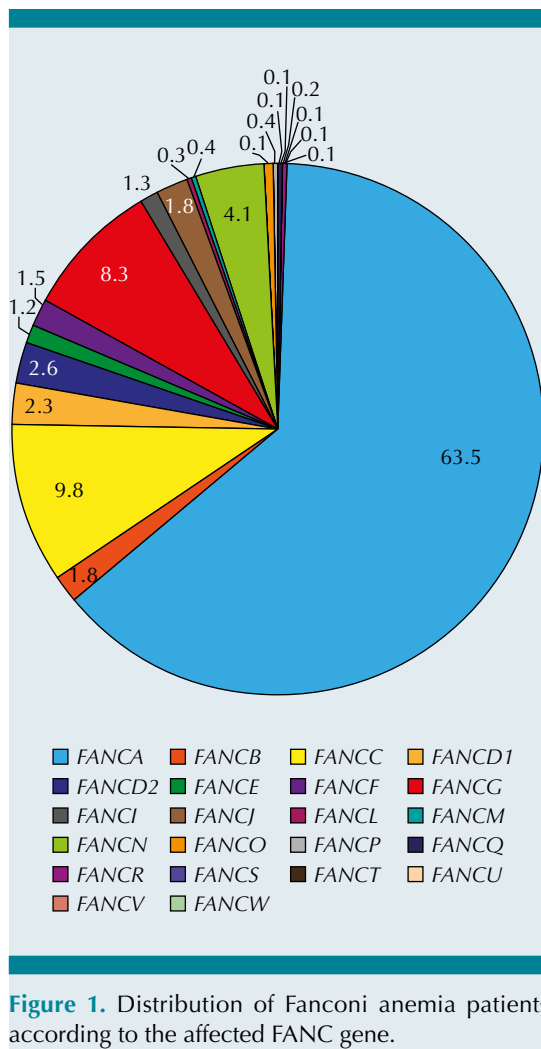
Table 1. Fanconi anemia genes and chromosomal location

Gene/Alias	Cytogenetic location
<i>FANCA</i>	16q24.3
<i>FANCB</i>	Xp22.2
<i>FANCC</i>	9q22.32
<i>FANCD1/BRCA2</i>	13q13.1
<i>FANCD2</i>	3p25.3
<i>FANCE</i>	6p21.31
<i>FANCF</i>	11p14.3
<i>FANCG/XRCC9</i>	9p13.3
<i>FANCI</i>	15q26.1
<i>FANCI/BRIP1</i>	17q23.2
<i>FANCL</i>	2p16.1
* <i>FANCM</i>	14q21.2
<i>FANCN/PALB2</i>	16q12.2
* <i>FANCO/RAD51C</i>	17q22
<i>FANCP/SLX4</i>	16p13.3
<i>FANCP/XPF</i>	16p13.12
* <i>FANCR/RAD51</i>	15q15.1
* <i>FANCS/BRCA1</i>	17q21.31
<i>FANCT/UBE2T</i>	1q32.1
<i>FANCU/XRCC2</i>	7q36.1
<i>FANCV/REV7</i>	1p36.22
<i>FANCW/RFWD3</i>	16q23.1

* These genes are called “Fanconi anemia-like” because patients who carry a PV in any of them, do not present bone marrow failure 1.

The fact that several genes and hundreds of unique PV are associated with the FA phenotype makes locus and allelic heterogeneity the canon in this disease. Because of this complexity, molecular diagnosis of the disease can be challenging.

Different working groups in the world have followed different strategies according to their population and available resources.⁴⁻⁶ Finding the appropriate molecular strategy to study a group of FA patients can be a long and costly task.



The objective of this communication is to present the strategy followed by our group to determine the genotype of 50 patients with FA, in 40 of whom (80%), we were able to obtain the causal genotype of the disease. We hope that this work contributes to finding the most economical and effective genotyping strategy for patients with FA, in laboratories with characteristics such as ours.

We start with biological samples from patients with a clinical and cytogenetic diagnosis of FA. The samples usually are obtained from peripheral

blood or bone marrow aspirate, from which we extract genomic DNA. Thereafter, we proceed according to the workflow suggested below (**Figure 2**):

1. We screen the *FANCA* gene through a Multiplex Ligation-Dependent Probe Amplification (MLPA) assay to detect long deletions, because *FANCA* deletions are one of the most frequent PV reported, and these deletions can range from one or multiple exons to the entire gene.
2. For negative patients, we perform MLPA analysis for *FANCD2*, *FANCN* (*PALB2*) and *FANCB* genes, which also present large deletions and duplications detectable by MLPA.
3. If no PV are found, or if they are detected in only one of the two alleles, we continue genotyping through next-generation sequencing (NGS), either by a targeted panel (in our case, composed of the 16 *FANCA* genes more commonly affected), or by whole-exome sequencing (WES), to detect small or point PV. In this way we found the variants responsible for the disease in most of the patients studied.
4. If we still have patients without genotype, the screening method we have followed is high-resolution microarrays with genome-wide markers for single nucleotide polymorphism (SNP) and copy number variants (CNV), for the search of large deletions or duplications at the *FANCA* genes. Microarrays also allow the detection of regions with long contiguous stretches of homozygosity (LCSH), in order to find uninterrupted homozygous alleles with neutral copy number state of 2 at the genomic level.

LCSH are observed throughout the human genome due to endogamy (consanguinity or reproductive isolation, both geographic and

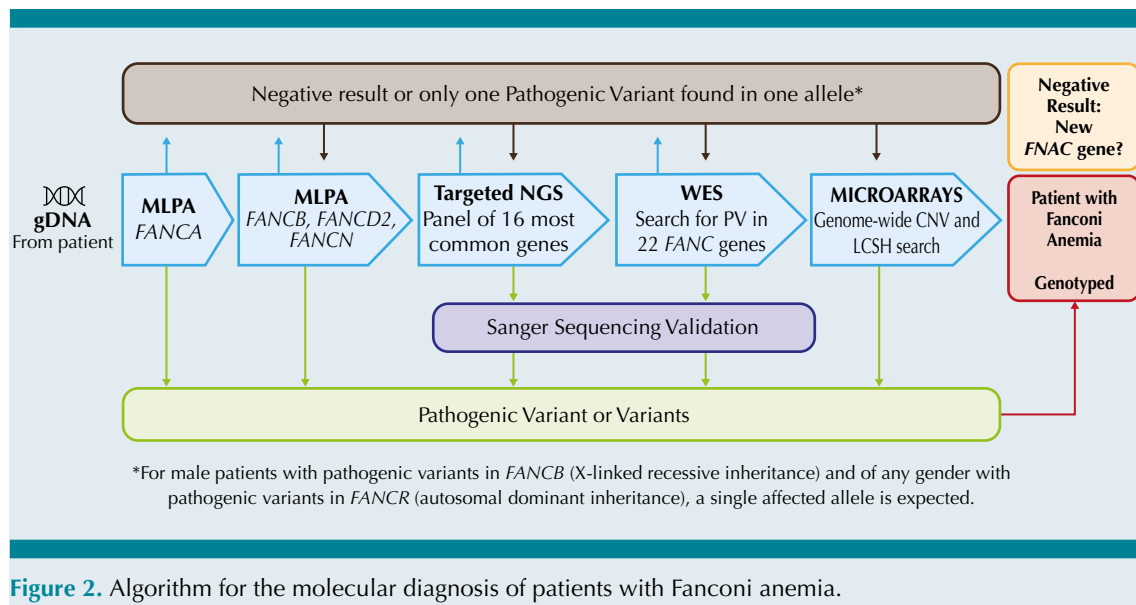


Figure 2. Algorithm for the molecular diagnosis of patients with Fanconi anemia.

social) or evolutionary forces. Although LCSH are typical for inbred populations, they are also common and sometimes unexpectedly long in the genome of outbred people.⁷⁻¹⁰ The importance of their detection is that they could indicate the presence of genes in a recessive state and if they are detected in an area that contains any of the 20 FANC genes with autosomal recessive inheritance, an intentional search can be made to detect VP in that locus (**Figure 2**).

Methodology

Patients

We include patients with clinical features of FA who tested positive for cytogenetic chromosome breakage. We explain the procedure and ask for written informed consent before obtaining the biological samples.

Genomic DNA extraction

Blood or bone marrow aspirate should be collected in tubes containing EDTA as an anticoagulant. One of the most frequent clinical features in FA patients is the low count of white

cells in blood. Consequently, we expect to have a small starting material for the genomic DNA (gDNA) extraction.

We use the Gentra Puregene Blood Kit (QIAGEN, cat 158389, Venlo, Limburg, NL) (https://www.qiagen.com/Gentra_Puregene_Handbook_12/2014_33) with a few modifications. This protocol purifies gDNA from fresh or frozen samples of 3 ml whole blood. The protocol can also be used for gDNA purification from packed cells, buffy coat, or bone marrow. If processing packed cells or buffy coat samples, directly scale the volume of reagents used in proportion to the volume of the original blood sample.

Procedure

- a. Dispense 3 ml whole blood or bone marrow in a 15 ml centrifuge tube, add 9 ml Red Blood Cell (RBC) Lysis Solution, and mix by inverting 10 times.
- b. Incubate 10 min at room temperature (15–25°C). Invert at least once during the incubation.

- c. Centrifuge for 5 min at 2000 Relative Centrifugal Force (RCF) to pellet the white blood cells; if there are remaining red cells, repeat the RBC lysis step.
- d. Discard the supernatant carefully by pipetting or pouring, leaving approximately 200 μ l of the residual liquid and the white blood cell pellet. Vortex the tube vigorously to resuspend the pellet in the residual liquid, the pellet should be completely dispersed after vortexing.
- e. Add 3 ml Cell Lysis Solution, and pipet up and down to lyse the cells or vortex vigorously for 10 s and incubate at 37°C until the solution is homogeneous.
- f. Add 15 μ l RNase A [4 mg/ml] Solution and mix by inverting 25 times. Incubate for 30 min at 37°C. Then add 15 μ l Proteinase K [20mg/ml] and incubate for 30 min at 55°C. Immediately put in ice to cool the sample.
- g. Add 1 ml Protein Precipitation Solution, vortex vigorously for 20 s at high speed and incubate on ice for 15 min.
- h. Centrifuge for 10 min/4°C at 2000 RCF. The precipitated proteins should form a tight, dark brown pellet. If the protein pellet is not tight, repeat the centrifugation.
- i. Dispense the supernatant from the previous step in a new 15 ml centrifuge tube. Be sure the protein pellet is not dislodged during pouring. Add 3 ml of isopropanol and 1 μ l of glycogen [20 mg/ml]
- j. Mix by inverting gently 50 times until the gDNA is visible as threads or a clump. Centrifuge for 10 min at 2000 RCF. The gDNA may be visible as a small white pellet.
- k. Carefully discard the supernatant or remove with a pipette tip. Add 3 ml of 70% ethanol and invert several times to wash the gDNA pellet.
- l. Centrifuge for 5 min at 2000 x RCF. Remove the supernatant, transfer the pellet with the remaining liquid to a 1.5 ml microcentrifuge tube, and wash again with 300 μ l of 70% ethanol. Centrifuge, remove the supernatant, and let dry the DNA in an open tube at room temperature or 65°C.
- m. Resuspend the gDNA pellet in 50 μ l of nuclease-free low Tris-EDTA (TE) buffer or nuclease-free water.
- n. The gDNA obtained can be quantified by a Nanodrop™ spectrophotometer (Nanodrop Technologies, Wilmington, DE). For NGS assays a more sensitive quantification could be required, such as fluorometric quantification (Qubit, Thermofisher).
- o. The integrity of the gDNA is evaluated by standard agarose gel electrophoresis 1% in Tris Acetate EDTA (TAE) 1X buffer.

Multiplex Ligation-Dependent Probe Amplification (MLPA)

The MLPA is a multiplex PCR-based assay for detecting CNV. In this method, probes specific to a genomic region hybridize to gDNA and then are amplified by PCR using a single pair of oligonucleotides.¹¹ The great advantage of this method is that it allows the simultaneous and semi-quantitative amplification of more than 60 probes in a single reaction for the rapid and accurate identification of deletions and/or duplications.¹²

PV found in *FANCA* accounts for approximately 70% of reported cases, of which 15% to 20% correspond to large deletions.^{13,14} Most of them have been found mainly through MLPA, often spanning multiple exons.¹⁵ We search these alterations with the MLPA kit developed by MRC Holland (<https://www.mrcholland.com/product/P057>).

In addition to *FANCA*, there are other *FANC* genes in which CNV have been frequently reported and that can also be assessed with this methodology: *FANCD2* (~ 3% of patients),^{16,17} *FANCN* (<1% of patients)¹⁸ and *FANCB* (~ 2% of patients).¹⁹

In our experience, the optimal amount of gDNA to be used for each reaction should be between 80-90 ng, diluted in a low TE buffer. On the other hand, each probe mix contains fragments that are used for quality control (usually 9), to evaluate parameters such as denaturation (“D” fragments), the amount of DNA, and the efficiency of the ligation step (“Q” fragments). It is important to include reference DNA samples from a) healthy individuals (both male and female), b) a FA positive sample (e.g., we use the FA cell line FA-A-VU817) and c) a negative sample (low TE buffer without DNA); these controls should be included during each run. It is suggested to use 3 reference samples for every 21 samples.

We follow the same workflow suggested in the general protocol by MRC Holland (<https://support.mrcholland.com/downloads/files/mlpa-general-protocol-one-tube>), which can be summarized in the following steps:

Procedure

- a. Denaturation of DNA and probe hybridization: Heat the DNA sample of interest (5 µl) for 5 minutes at 98°C. Then, add 3 µl of the probe master mix (1.5 µl SALSA probe mix + 1.5 µl MLPA buffer) to the DNA sample. Place the tube for 1 minute at 95°C, and hybridize for 16 h at 60°C.
- b. Ligation: After the hybridization time has elapsed, decrease the temperature of the thermal cycler to 54°C. Without removing the tubes from the equipment, add 32 µl of the ligation master mix and incubate at the same temperature for 15 min. To inactivate Ligase 65, incubate for 5 minutes at 98°C.
- c. PCR amplification: The reaction tubes are cooled to room temperature and 10 µl of the polymerase master mix is added. Start the PCR reaction, using the suggested program (95°C for 30 s, 60°C for 30 s and 72°C for 1 minute, for 35 cycles), with a final amplification at 72°C for 20 minutes. Store the tubes protected from light until capillary electrophoresis.
- d. Capillary electrophoresis: PCR products are loaded into a capillary electrophoresis instrument to be separated according to their length. The ratio of the components of the reaction mixture (final product of the MLPA reaction, weight marker, and formamide), as well as the electrophoretic run parameters, vary depending on the instrument available.
- e. Data analysis: Data review must be performed using the Coffalyzer program, developed by MRC Holland (<https://www.mrcholland.com/technology/software/coffalyser-net>). During this step, the relative copy number is determined by contrasting the relative peak heights of the reference and target probes in the test samples against those of the reference samples. It is essential to set up the study of the raw signals obtained with the version of each probe mix and the batch number with which the reaction was performed.

Once the electropherogram is obtained, the height of the fluorescence peaks after fragmentation is evaluated. The fluorescence level of each peak in a sample is compared with the signal observed in the DNA samples used as reference. Thus, deletion will be visualized as a decrease in peak height compared with the wild-type fragment, while a duplication increases the size of that fragment. Subsequently, the raw signals are normalized to obtain a value that reflects the adjusted ratio between the contrasting signals. This value is known as the dose coefficient

(DQ) and is used to discriminate the number of copies. The DQ ranges from 0 (homozygous deletions) to 2.15 (homozygous duplications). Values between 0.8 and 1 are considered normal (diploid set).²⁰

Next-Generation Sequencing (NGS)

NGS is a generic term used to describe the use of various technologies whose common feature is the massively parallel sequencing of multiple DNA fragments.²¹ The combination of high throughput and single DNA molecule sequencing are hallmarks of NGS, regardless of the sequencing platform used.²² The evolution and diversification of NGS assays have made it possible to overcome the limitations of Sanger sequencing and generate a huge amount of data, making it dependent on intensive computational analysis for the assembly of the sequences produced.²³

The main steps of DNA NGS involve DNA fragmentation, library preparation, sequencing, bioinformatics analysis, annotation of the variants found, and interpretation.²⁴ However, it should be noted that the choice of the desired fragment length (short vs. long “reads”) depends on the research question and its clinical application.

The application of NGS technologies has been especially useful in disorders with high phenotypic variability, accelerating the identification and discovery of new causal variants, which has boosted its application in molecular diagnostics.²⁵ The reduction of costs and the potential of the information obtained have allowed the integration of NGS data into the clinical practice of medicine.²⁶ Currently, there are different approaches to NGS, depending on whether the scope of sequencing involves a large number of genes (targeted gene panels), protein-coding regions (whole-exome sequencing (WES)), or the entire genome (whole-genome sequencing (WGS)).

According to the most recent version of the AF Clinical Care Guideline, NGS should be performed as part of the genotyping process, selecting a panel of *FANC* genes in the first stage, appealing to WES in case of negative results by panel test.²⁷ We have followed this strategy for most of the cases analyzed, considering the benefits and limitations of each approach.

The use of dedicated panels has become widespread in the FA context, because the better characterization of the FA/BRCA pathway has allowed studying more *FANC* genes in a single assay, in addition to offering greater depth of coverage at a relatively low cost. However, WES has become increasingly attractive and affordable due to its capacity to sequence genetic regions in candidate or novel *FANC* genes.²⁷ An example that illustrates this idea is given by the description of bi-allelic PV in a group of patients with FA without genotype classification in *ERCC4/XPF*, whose loss of function had previously been associated with Xeroderma pigmentosum. This finding led to this gene being reclassified as *FANCC*.²⁸ Similarly, PV were discovered in *FANCT* (also known as *UBE2T*, in two Japanese patients)²⁹ and *FANCU*.³⁰

Our search for PV in FA by NGS is summarized in the following bioinformatics workflow:

Procedure

- a. Gene panel design. The in-silico design of specific probes for 16 *FANC* genes (*FANCA*, *FANCB*, *FANCC*, *FANCD1/BRCA2*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *FANCI/BRIP1*, *FANCL*, *FANCM*, *FANCN/PALB2*, *FANCO/RAD51C*, *FANCP/SLX4*, and *FANCCQ/ERCC4*) is performed with the SureDesign web-based application (Agilent Technologies). The final design is expected to cover 99.5% of all coding exons of these genes plus 50 bp of their flanking 5' and 3' intronic regions.

- b.** Digestion, library synthesis, and PCR amplification of targeted fragments are achieved using the Enrichment Haloplex System (Agilent Technologies Inc., 2013, Waghäusel-Wiesental, Germany).
- c.** Short paired-end reads are obtained through sequencing by synthesis technology, on an Illumina MiSeq platform (Illumina, San Diego, CA, USA).
- d.** In case the bioinformatics analysis of the panel data is negative, we proceed to use the WES technique following the suggested methodological adaptations.

Unlike panel design, in the case of WES, there are different commercial options focused on the capture of exonic regions (e.g., SureSelectXT Human All Exon V4 (Agilent Technologies), KAPA HyperExome (Roche), among others). In addition, the use of reagents and kit formulations compatible with the aforementioned systems extends to the enrichment and library construction steps. Finally, the sequencing platforms are adapted for a more scalable throughput (e.g., NovaSeq 6000 system, Illumina), due to the amplification of a larger number of DNA targets.

Note that the technical and interpretative complexity of NGS usually restricts its use to highly qualified public or private laboratories which can accomplish the recommendations and requirements suggested in international standards.³¹

Because of these demands, these laboratories frequently offer NGS services to researchers and clinicians. We have used this modality to obtain sequencing data from the submission of high-quality DNA samples. The selection of the approach, as well as the platforms, is dependent on the project purpose and resources available for screening FA patients.

NGS data analysis

Due to the large amounts of data generated from NGS, even from small gene panels, computa-

tional analyses are required to effectively process and evaluate the sequence information.²⁵ The set of bioinformatics algorithms that are executed in a succession of defined steps to process the massive raw data generated from the NGS platform is known as the bioinformatics pipeline.²¹ Despite the heterogeneity of platforms and the scope of the study, the bioinformatics workflow consists of five major steps: quality control of raw data (from a text-based file format that contains sequence reads and individual base quality scores, known as FASTQ), preprocessing, mapping and postalignment, variant calling, and annotation and prioritization.^{23,32,33}

We developed an analytical pipeline (validated for data analysis) to streamline computational analysis and maximize variant calling accuracy, based on GATK best practice recommendations.³⁴⁻³⁶ The programs used for file manipulation are often packaged in the GATK Toolkit, a collection of command-line tools for analyzing high-throughput sequencing data with a primary focus on variant discovery. It is very important to highlight that these tools only run in Linux or MacOSX operating systems.

Procedure (the program's options are named in parentheses):

- a.** Review quality control of sequence reads (FastQC).
- b.** Filter and trim the reads with base poor quality and sequencing adapters (Trimomatic, Cutadapt, and Trim Galore).
- c.** Trimmed reads are aligned to map with a reference sequence (Burrows-Wheeler Alignment (BWA-MEM), using the references found at UCSC or Ensembl databases).
- d.** SAM files are transformed to BAM files and mark duplicates (Samtools and Picard).
- e.** Check the quality of BAM files (Qualimap).

f. Follow the recommendations of the GATK to get the variant calling file (VCF), in gVCF mode.

g. Finally, we annotate the variants (VEP, SnpEff, or ANNOVAR). Variants are described following the guidelines proposed by the Human Genome Variation Society (HGVS) nomenclature.³⁷

Variant Classification

Once we have filtered the candidate variants for each patient, the next step is to classify them according to their degree of pathogenicity. The most widely used system to perform this task comes from the recommendations suggested by the American College of Medical Genetics and the Association for Molecular Pathology (ACMG-AMP).^{38,39} According to this guide, the criteria for variant classification is a score based on five classes: benign, likely benign, variant of unknown significance (VUS), likely pathogenic, and pathogenic.

The criteria used for this classification are based on an evidence framework, in which different parameters (functional, population, segregation data, computational and predictive algorithms, among others) are considered to establish a final statement. Several web tools facilitate the classification exercise, as they incorporate valuable information from databases (such as ExAc, ClinVar, gnomAD, among others) into a single platform (e.g. Varsome).

NGS validation

Because NGS is not an error-free technique, the ACMG and the European Society of Human Genetics (ESHG) strongly recommend validating the selected PV. Technical as well as bioinformatics factors are related to the difficulty of confirming variants, mainly in genes with pseudogenes, repetitive regions, and with high GC content or hotspot variants. Sanger sequencing can solve this problem.

In our case, we design specific flanking primer pairs to validate the PV founded by NGS, using the Primer3 algorithm (<http://bioinfo.ut.ee/primer3-0.4.0/>), one of the most widely used primer designing tools⁴⁰. PCR amplification is carried out from 100 ng of DNA template. PCR products are purified with QIAquick PCR Purification kit (QIAGEN, Venlo, Limburg, NL) according to manufacturer instructions. Purified amplicons are bidirectional sequenced using the Big Dye Terminator sequencing kit and resolved on an Applied Biosystems 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

The electropherograms are analyzed with Chromas V2.6.6 software (www.technelysium.com.au) and target sequences are compared to the corresponding reference sequences from GenBank (<https://www.ncbi.nlm.nih.gov>). There are other alternative programs for analysis (<https://www.genewiz.com/Public/Resources/Tools-for-Viewing-Sequencing-Data/>) or web applications such as GEARgenomics (<https://www.gear-genomics.com/>), supported by EMBL-EBI, EMBL-European Bioinformatics Institute).

Sanger sequencing as genotyping strategy

Sanger sequencing could be performed to search PV in specific genes, i.e., when members of a family with a previously identified PV are studied or in individuals belonging to a community where a founder mutation has been reported, particularly, in genetically isolated populations such as the FANCA c.295C>T in Spanish Gypsies,⁴¹ FANCC c.456+4A>T in Ashkenazi Jews,⁴² FANCC c.67delG in the Mennonite community⁴³ and FANCG c.511-3_511-2delCA in a Mexican Mixe community.⁴⁴

High-resolution microarray analysis

We perform the analysis of the gDNA with Affymetrix CytoScan750K arrays (Affymetrix, Santa Clara, CA, USA). This array provides a genome-wide coverage, including 550,000 markers for

detecting copy number variation and 200,436 SNP probes.

The steps for DNA digestion, ligation, PCR amplification, fragmentation, labeling, denaturing, and hybridization into the array are performed according to the protocols and QC guidelines provided by the supplier.

Arrays are then stained and washed in the Affymetrix GeneChip Fluidics Station 450 and scanned using an Affymetrix GeneChip Scanner 3000 7G (Affymetrix, Santa Clara, CA, USA).

The files obtained were analyzed with the software Chromosome Analysis Suite (ChAS) software version 4.1, provided by Affymetrix (Affymetrix, Santa Clara, CA, USA). (NetAffx 33 hg19 annotation files) (<http://www.affymetrix.com>). The reporting threshold of the copy number (CN) is performed at "high resolution" settings: gains or losses in 100Kb with marker count ≥ 50 probes.

Analysis of long-contiguous stretch of homozygosity (LCSH)

Minimal thresholds for LCSH calls in clinical analyses are generally set at 3 to 10 Mb.⁴⁵ We visualize the LCSH in ChAS software, version 4.1, provided by Affymetrix (Affymetrix, Santa Clara, CA, USA). For the analysis, the NetAffx 33 hg19 annotation files are used, (<http://www.affymetrix.com>). For LCSH >3 Mb, the analysis configuration is set at LOH with marker count = 50 and size = 3000 Kbp, and for LCSH >5 Mb is set at marker count = 50 and size = 5000 Kbp.

Results report

Once the variants have been classified and validated, the next step is to write a report. In addition to the ACMG, other organizations, such as Clinical Genome Resource (ClinGen), European Society of Human Genetics (ESHG), and the Association for Clinical Genetic Science

(ACGS) have developed guidelines to advise laboratories in reporting and describing variants^{38,39,46-48}. These documents are issued from a laboratory to a treating physician and describe the results of genetic analysis. They should be clear, concise, and follow the recommendations suggested in existing guidelines⁴⁸.

Recommendations

- a. In general, the report format includes information on the requesting clinician, the laboratory where the analysis was performed, and the patient.
- b. For the description and interpretation of the results, the nomenclature of the variants should be written according to the Human Genome Organization (HUGO, <https://www.genenames.org/>) and the Human Genome Variation Society (HGVS, <https://varnomen.hgvs.org/>).

The report should include a clear interpretation of the test results based on a review of the patient's clinical context and available family information.⁴⁸ The goal should be to generate a simple, unambiguous statement that explains the clinical implications of the result.

- c. Methodological details (principle of the technology used, sensitivity, resolution, limitations, bioinformatics tools, among others) relevant to the interpretation of the results should be included in the report. For example, in the case of massive sequencing data for constitutional variants, it is advisable to refer to the most recent guidelines developed by the ACMG.
- d. Each laboratory should document the evidence used to classify variants concerning the impact they may have on the condition studied and establish precise policies to define which variants will be excluded from reporting.³⁹

In 2008, the Fanconi Anemia Mutation Database was created as a cooperative effort to accelerate access to information on variants found in the *FANC* genes. These data are available using the Leiden Open-Source Variation Database (<https://www2.rockefeller.edu/fanconi/mutate/>). The addition of a variant (new or previously reported) should be done as an institution and not on an individual basis. Sharing data with the scientific community significantly increases the amount of existing evidence for a particular variant in public databases and contributes to improving approaches in the search for genotype-phenotype correlations.⁴⁹ Additionally, this practice should be encouraged because it facilitates variant reclassification, which is especially important in the case of VUS, where the use of this information adds to the evidence generated through predictive biochemical algorithms and functional assay validation.⁵⁰

- e. Negative results in a patient with positive chromosomal breakage study.

Negative molecular test results should be carefully interpreted for an individual with a chromosomal breakage test within the FA range. One explanation for a negative result is the presence of a variant(s) in an undiscovered *FANC* gene or a type of variant in a gene that cannot be identified with the technology employed; another possibility is somatic mosaicism.^{51,52} Analysis of a tissue alternative to the peripheral blood or bone marrow, such as fibroblasts, may be considered in individuals presenting with an FA phenotype and negative genetic studies on peripheral blood.

DISCUSSION

In a variable percentage of patients (in our case it was 20%), it is not possible to reach the genotype. Possible reasons are that they could be variants that are not detected by exome sequencing, such as those found in intronic or regulatory regions of genes, and that would

require another more expensive strategy, whole genome sequencing. Alternatively, it could be because the variants are found in genes not yet described as *FANC* genes, so it is advisable to analyze the genes functionally associated with the FA/BRCA pathway.

The great genetic heterogeneity of FA makes it difficult to implement detection techniques classically used in other types of diseases of Mendelian inheritance. For this reason, we consider that the detailed disclosure of the methodological aspects inherent to genotyping in FA is essential to establish the minimum biological and technological requirements to identify the variants responsible for the disease, especially in geographical areas where this approach is not used or is limited. This methodological strategy, which ranges from obtaining the gDNA to the generation of the report, puts in context the importance of the information obtained (such as the frequency, location and type of variant) to develop a cost-effective and appropriate genotyping strategy for the FA population.

CONCLUSIONS

Describing the FA genotype in a population is relevant for several reasons: first, it improves the characterization of patients and thus contributes to the identification of FA genotype-phenotype correlations. Second, it refines the tailoring of genetic diagnostic strategies in specific groups of patients, including adequate selections of non-FA sibling bone marrow donors and appropriate genetic counseling. Finally, the description of further cases with seldomly described genotypes brings new data to consider when confronted with variants with a history of ambiguous pathogenicity classification

Acknowledgements

This research was partially funded by Consejo Nacional de Ciencia y Tecnología (CONACyT), grant number SALUD-2014-1-233721; Insti-

tuto Nacional de Pediatría, Recursos Fiscales E022 project INP 2014/41; Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), Universidad Nacional Autónoma de México (UNAM) Grant Number IN205120. The authors thank M. en C. Francisco Zúñiga-Frias for the English revision of this manuscript.

REFERENCIAS

- Fiesco-Roa MÓ, García-de Teresa B, Leal-Anaya P, van 't Hek R, Wegman-Ostrosky T, Frías S, et al. Fanconi anemia and dyskeratosis congenita/telomere biology disorders: Two inherited bone marrow failure syndromes with genomic instability. *Front Oncol* [Internet]. 2022;12. Available from: <https://www.frontiersin.org/articles/10.3389/fonc.2022.949435>
- García-de-Teresa B, Rodríguez A, Frías S. Chromosome Instability in Fanconi Anemia: From Breaks to Phenotypic Consequences. *Genes (Basel)*. 2020 Dec;11(12).
- Niraj J, Färkkilä A, D'Andrea AD. The Fanconi Anemia Pathway in Cancer. *Annu Rev cancer Biol*. 2019 Mar;3:457–78.
- Pilonetto D V, Pereira NF, Bonfim CMS, Ribeiro LL, Bitencourt MA, Kerkhoven L, et al. A strategy for molecular diagnostics of Fanconi anemia in Brazilian patients. *Mol Genet genomic Med*. 2017 Jul;5(4):360–72.
- Mori M, Hira A, Yoshida K, Muramatsu H, Okuno Y, Shiraishi Y, et al. Pathogenic mutations identified by a multimodality approach in 117 Japanese Fanconi anemia patients. *Haematologica*. 2019 Oct;104(10):1962–73.
- Ben Haj Ali A, Messaoud O, Elouej S, Talmoudi F, Ayed W, Mellouli F, et al. FANCA Gene Mutations in North African Fanconi Anemia Patients. *Front Genet* [Internet]. 2021;12. Available from: <https://www.frontiersin.org/articles/10.3389/fgene.2021.610050>
- Li LH, Ho SF, Chen CH, Wei CY, Wong WC, Li LY, et al. Long contiguous stretches of homozygosity in the human genome. *Hum Mutat*. 2006 Nov;27(11):1115–21.
- Pajusalu S, Žilina O, Yakoreva M, Tammur P, Kuuse K, Mõlter-Väär T, et al. The Diagnostic Utility of Single Long Contiguous Stretches of Homozygosity in Patients without Parental Consanguinity. *Mol Syndromol*. 2015 Sep;6(3):135–40.
- Wang JC, Ross L, Mahon LW, Owen R, Hemmat M, Wang BT, et al. Regions of homozygosity identified by oligonucleotide SNP arrays: evaluating the incidence and clinical utility. *Eur J Hum Genet*. 2015 May;23(5):663–71.
- Chaves TF, Oliveira LF, Ocampos M, Barbato IT, de Luca GR, Barbato Filho JH, et al. Long contiguous stretches of homozygosity detected by chromosomal microarrays (CMA) in patients with neurodevelopmental disorders in the South of Brazil. *BMC Med Genomics*. 2019 Mar;12(1):50.
- Stuppia L, Antonucci I, Palka G, Gatta V. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. *Int J Mol Sci*. 2012;13(3):3245–76.
- Hömig-Hölzel C, Savola S. Multiplex ligation-dependent probe amplification (MLPA) in tumor diagnostics and prognostics. *Diagnostic Mol Pathol Am J Surg Pathol part B*. 2012 Dec;21(4):189–206.
- Ameziane N, Errami A, Léveillé F, Fontaine C, de Vries Y, van Spaendonk RML, et al. Genetic subtyping of Fanconi anemia by comprehensive mutation screening. *Hum Mutat*. 2008 Jan;29(1):159–66.
- Castella M, Pujol R, Callén E, Trujillo JP, Casado JA, Gille H, et al. Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations. *Blood*. 2011 Apr;117(14):3759–69.
- Kimble DC, Lach FP, Gregg SQ, Donovan FX, Flynn EK, Kamat A, et al. A comprehensive approach to identification of pathogenic FANCA variants in Fanconi anemia patients and their families. *Hum Mutat*. 2018 Feb;39(2):237–54.
- Kalb R, Neveling K, Hoehn H, Schneider H, Linka Y, Batish SD, 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 May;80(5):895–910.
- Flynn EK, Kamat A, Lach FP, Donovan FX, Kimble DC, Narisu N, et al. Comprehensive analysis of pathogenic deletion variants in Fanconi anemia genes. *Hum Mutat*. 2014 Nov;35(11):1342–53.
- Xia B, Dorsman JC, Ameziane N, de Vries Y, Roomians MA, Sheng Q, et al. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat Genet*. 2007 Feb;39(2):159–61.
- Jung M, Ramanagoudr-Bhojappa R, van Twest S, Rosti RO, Murphy V, Tan W, et al. Association of clinical severity with FANCB variant type in Fanconi anemia. *Blood*. 2020 Apr;135(18):1588–602.
- Coffa J, van den Berg J. Analysis of MLPA Data Using Novel Software Coffalyser.NET by MRC-Holland. In: Eldin AB, editor. *Modern Approaches To Quality Control* [Internet]. Rijeka: IntechOpen; 2011. Available from: <https://doi.org/10.5772/21898>
- Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, et al. Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines: A Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J Mol Diagn*. 2018 Jan;20(1):4–27.
- Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. *Hum Immunol*. 2021 Nov;82(11):801–11.
- Voelkerding K V, Coonrod EM, Durtschi JD, Margraf RL. Next-Generation Sequencing: Principles for Clinical Application BT - Molecular Pathology in Clinical Practice. In: Leonard DGB, editor. Cham: Springer International Publishing; 2016. p. 889–909. Available from: https://doi.org/10.1007/978-3-319-19674-9_59

24. Qin D. Next-generation sequencing and its clinical application. *Cancer Biol Med*. 2019 Feb;16(1):4–10.
25. den Dunnen JT. Efficient variant data preparation for Human Mutation manuscripts: Variants and phenotypes. Vol. 40, Human mutation. United States; 2019. p. 1009.
26. Rizzo JM, Buck MJ. Key principles and clinical applications of “next-generation” DNA sequencing. *Cancer Prev Res (Phila)*. 2012 Jul;5(7):887–900.
27. Sroka I, Frohnmayer L, Wirkkula L. Fanconi Anemia Clinical Care Guidelines [Internet]. Fifth Edit. Vol. 1, Fanconi Anemia Research Fund, Inc. 2020. 1–276 p. Available from: <http://marefateadyan.nashriyat.ir/node/150>
28. Bogliolo M, Schuster B, Stoeper C, Derkunt B, Su Y, Raams A, et al. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet*. 2013 May;92(5):800–6.
29. Hira A, Yoshida K, Sato K, Okuno Y, Shiraishi Y, Chiba K, et al. Mutations in the gene encoding the E2 conjugating enzyme UBE2T cause Fanconi anemia. *Am J Hum Genet*. 2015 Jun;96(6):1001–7.
30. Shamseldin HE, Elfaki M, Alkuraya FS. Exome sequencing reveals a novel Fanconi group defined by XRCC2 mutation. *J Med Genet*. 2012 Mar;49(3):184–6.
31. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med Off J Am Coll Med Genet*. 2013 Sep;15(9):733–47.
32. McCombie WR, McPherson JD, Mardis ER. Next-Generation Sequencing Technologies. *Cold Spring Harb Perspect Med*. 2019 Nov;9(11).
33. Bao R, Huang L, Andrade J, Tan W, Kibbe WA, Jiang H, et al. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Inform*. 2014;13(Suppl 2):67–82.
34. DePristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011 May;43(5):491–8.
35. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma*. 2013;43(1110):11.10.1-11.10.33.
36. Franke KR, Crowley EL. Accelerating next generation sequencing data analysis: an evaluation of optimized best practices for Genome Analysis Toolkit algorithms. *Genomics Inform*. 2020 Mar;18(1):e10.
37. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum Mutat*. 2016 Jun;37(6):564–9.
38. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med Off J Am Coll Med Genet*. 2015 May;17(5):405–24.
39. Rehder C, Bean LJH, Bick D, Chao E, Chung W, Das S, et al. Next-generation sequencing for constitutional variants in the clinical laboratory, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med Off J Am Coll Med Genet*. 2021 Aug;23(8):1399–415.
40. Kumar A, Chordia N. In Silico PCR Primer Designing and Validation BT - PCR Primer Design. In: Basu C, editor. New York, NY: Springer New York; 2015. p. 143–51. Available from: https://doi.org/10.1007/978-1-4939-2365-6_10
41. Callén E, Casado JA, Tischkowitz MD, Bueren JA, Creus A, Marcos R, et al. A common founder mutation in FANCA underlies the world’s highest prevalence of Fanconi anemia in Gypsy families from Spain. *Blood*. 2005 Mar;105(5):1946–9.
42. Whitney MA, Jakobs P, Kaback M, Moses RE, Grompe M. The Ashkenazi Jewish Fanconi anemia mutation: incidence among patients and carrier frequency in the at-risk population. *Hum Mutat*. 1994;3(4):339–41.
43. García-de Teresa B, Frias S, Molina B, Villarreal MT, Rodríguez A, Carnevale A, et al. FANCC Dutch founder mutation in a Mennonite family from Tamaulipas, México. *Mol Genet genomic Med*. 2019 Jun;7(6):e710.
44. Reyes P, García-de Teresa B, Juárez U, Pérez-Villatoro F, Fiesco-Roa MO, Rodríguez A, et al. Fanconi Anemia Patients from an Indigenous Community in Mexico Carry a New Founder Pathogenic Variant in FANCG. *Int J Mol Sci*. 2022 Feb;23(4).
45. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med Off J Am Coll Med Genet*. 2011 Jul;13(7):680–5.
46. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, et al. ClinGen—the Clinical Genome Resource. *N Engl J Med*. 2015 Jun;372(23):2235–42.
47. Matthijs G, Souche E, Alders M, Corveleyn A, Eck S, Feenstra I, et al. Guidelines for diagnostic next-generation sequencing. *Eur J Hum Genet*. 2016 Oct;24(10):1515.
48. Smith K, Martindale J, Wallis Y, Bown N, Leo N, Creswell L, et al. General genetic laboratory reporting recommendations. *Birmingham Assoc Clin Genet Sci*. 2015;11.
49. Directors ABO. Laboratory and clinical genomic data sharing is crucial to improving genetic health care: a position statement of the American College of Medical Genetics and Genomics. *Genet Med Off J Am Coll Med Genet*. 2017 Jul;19(7):721–2.
50. Federici G, Soddu S. Variants of uncertain significance in the era of high-throughput genome sequencing: a lesson from breast and ovary cancers. *J Exp Clin Cancer Res*. 2020 Mar;39(1):46.



51. Gregory JJ, Wagner JE, Verlander PC, Levran O, Batish SD, Eide CR, et al. Somatic mosaicism in Fanconi anemia: evidence of genotypic reversion in lymphohematopoietic stem cells. *Proc Natl Acad Sci U S A*. 2001 Feb;98(5):2532–7.
52. Ikeda H, Matsushita M, Waisfisz Q, Kinoshita A, Oostra AB, Nieuwint AWM, et al. Genetic reversion in an acute myelogenous leukemia cell line from a Fanconi anemia patient with biallelic mutations in BRCA2. *Cancer Res*. 2003 May;63(10):2688–94.