



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Universitat Politècnica de València

Escuela Técnica Superior de Ingeniería Agronómica y del
Medio Natural (ETSIAMN)

Trabajo Final de Grado en Biotecnología

Diseño y validación de un panel de genes por hibridación y captura dirigido a Neoplasias Mieloides Hereditarias

Autor:

Paula Carbonell Pérez

Tutores:

Ángel Zúñiga Cabrera
Rafael Sirera Pérez

Curso académico 2018/2019

Valencia, julio de 2019

DISEÑO Y VALIDACIÓN DE UN PANEL DE GENES POR HIBRIDACIÓN Y CAPTURA DIRIGIDO A NEOPLASIAS MIELOIDES HEREDITARIAS

Autor: Paula Carbonell Pérez

Localidad y fecha: Valencia, julio de 2019

Tutores: Ángel Zúñiga Cabrera y Rafael Sirera Pérez

RESUMEN

Las Neoplasias Mieloides Hereditarias consisten en un grupo de trastornos hematológicos de considerable incidencia y con una demostrada base genética, encontrándose una predisposición a su desarrollo en la herencia de alteraciones en línea germinal de diversos genes. Es por ello que se propone el desarrollo como técnica de diagnóstico, de un panel de genes dirigido, con la finalidad de poder detectar aquellas variantes heredadas que predisponen al desarrollo de este tipo de patología. El empleo de paneles genéticos hoy día se considera de mayor rentabilidad diagnóstica que los exomas dirigidos, ya que permite realizar diseños más manejables y donde se puedan tener en cuenta las características propias de los genes estudiados.

Para el desarrollo del panel, se ha realizado un diseño *custom*, para lo cual se empezó por la realización de una revisión bibliográfica de los artículos relacionados con el tema para tratar de obtener la información más actual posible al respecto. Una vez seleccionados los genes diana, buscamos las regiones de interés de éstos -zonas exónicas, intrónicas situadas a +/- 10 nucleótidos, variantes de interés clínico situadas en regiones reguladoras... - a partir del manejo de la información extraída de diversas bases de datos, con el objetivo de acotar al máximo las regiones de interés, ya que el espacio que ofrecen los paneles *custom* es limitado.

Una vez completado el diseño, es necesario validar la primera versión del panel de genes dirigido a través de diversos análisis que permitirán la obtención de datos, con el objetivo de establecer su sensibilidad y especificidad, utilizando para ello líneas celulares Coriell con variantes detectadas y validadas mediante distintas aproximaciones de secuenciación masiva, de manera que se consideran como patrones para la técnica de NGS (*Next Generation Sequencing*). El proceso de validación del panel es necesario antes de su empleo como herramienta diagnóstica.

PALABRAS CLAVE: Neoplasias Mieloides Hereditarias, predisposición, línea germinal, panel de genes dirigido, NGS, captura.

ABSTRACT

Hereditary Myeloid Neoplasms consist of a group of hematological disorders of considerable incidence and with a proven genetic base, finding a predisposition to their development in the inheritance of germline alterations of several genes. That is why we propose the development as a diagnostic technique, a directed gene panel, in order to detect those inherited variants that predispose to the development of this type of pathology. The use of genetic panels today is considered to be of greater diagnostic value than directed exomes, since it allows for more manageable designs and where the characteristics of the studied genes can be taken into account.

For the panel development, a custom design has been made, for which a bibliographic review of the articles related to the subject was started in order to obtain the most up-to-date information possible. Once the target genes have been selected, we look for the regions of interest of these - exonic, intronic zones located at +/- 10 nucleotides, variants of clinical interest located in regulatory regions ... - from the handling of the information extracted from various databases , with the aim of limiting regions of interest to the maximum, since the space offered by custom panels is limited.

Once the design is completed, it is necessary to validate the first version of the directed gene panel through various tests that will allow obtaining data, in order to establish its sensitivity and specificity, using Coriell cell lines with detected and validated variants using different approaches of massive sequencing, so that they are considered as standards for the NGS technique (Next Generation Sequencing). The validation process of the panel is necessary before its use as a diagnostic tool.

KEYWORDS: Hereditary Myeloid Neoplasms, predisposition, germline, targeted gene panel, NGS, capture.

AGRADECIMIENTOS

Qué bonito poder disponer de un apartado para agradecer el apoyo de todos aquellos que directa o indirectamente han participado en este trabajo, ya sea tanto a nivel profesional como a nivel personal, porque nunca deberíamos olvidar que, al fin y al cabo, somos personas.

Quería empezar por agradecer por su buen trato desde el primer día a José Miguel Molina, que con tanto cariño me ha aconsejado siempre desde el día que entré por primera vez en el Instituto de Investigación Sanitaria La Fe, cuando realicé mis primeras prácticas. Gracias por todas las horas extra de reflexión, orientación y consejos, pues gracias a ti he podido tomar mejores decisiones y encontré también esta posibilidad de realizar mi Trabajo Final de Grado.

También quería agradecer a Ángel Zúñiga la oportunidad que me ofreció de llevar a cabo este proyecto con él, convirtiéndose en mi cotutor académico, quien ha colaborado en su desarrollo dirigiéndome en cada uno de los pasos que he ido realizando a lo largo de estos meses. Y también me gustaría dar las gracias por su colaboración a Rafael Sirera, mi tutor académico durante el desarrollo de mi Trabajo Final de Grado.

Además, creo que es importante hacer un pequeño homenaje a todos aquellos profesores que a lo largo de la carrera han conseguido motivarnos, enseñarnos y mantener nuestro interés por la ciencia en un ambiente que tan competitivo se presenta actualmente para el biotecnólogo egresado.

Por último, quería dedicar unas palabras con mucho cariño a todos aquellos que forman parte de mi vida, a todos los que están ahí, desinteresadamente. Gracias, familia y amigos, lo sois todo para mí.

Al final, todo esfuerzo tiene su recompensa.

Índice

1. INTRODUCCIÓN	1
1.1. Neoplasias Mieloides Hereditarias	1
1.1.1. Clasificación de la OMS de las Neoplasias Mieloides Hereditarias	3
1.1.1.1. Neoplasias Mieloides con mutaciones predisponentes en línea germinal sin desórdenes plaquetarios o disfunción de órganos.....	4
1.1.1.2. Neoplasias Mieloides con mutaciones predisponentes de la línea germinal y desórdenes plaquetarios preexistentes	5
1.1.1.3. Neoplasias Mieloides con mutaciones predisponentes de la línea germinal y disfunción de órganos	6
1.1.2. Diagnóstico y manejo de pacientes con Neoplasias Mieloides Hereditarias.....	7
1.2. Uso de paneles de genes dirigidos para detectar mutaciones	9
2. HIPÓTESIS	11
3. OBJETIVOS	12
4. MATERIALES Y MÉTODOS.....	13
4.1. Diseño del panel de genes dirigido	13
4.1.1. Búsqueda bibliográfica.....	13
4.1.2. Recopilación de información de bases de datos	13
4.1.3. Diseño final del panel de genes	13
4.2. Validación del panel de genes dirigido	14
4.2.1. Preparación del ADN de las líneas celulares.....	14
4.2.2. Preparación de la librería enriquecida.....	15
4.2.2.1. Preparación de la librería de DNA genómico	15
4.2.2.2. Hibridación.....	17
4.2.2.3. Indexado	17
4.2.3. Secuenciación multiplex <i>paired-end</i>	18
4.2.4. Análisis bioinformático	18
5. RESULTADOS Y DISCUSIÓN.....	20

5.1.	Diseño obtenido del panel de genes dirigido	20
5.2.	Resultados de la validación del panel de genes dirigido	24
5.2.1.	Cobertura	24
5.2.2.	Precisión.....	25
5.2.2.1.	Línea celular Coriell D1 (NA12144).....	25
5.2.2.2.	Línea Celular Coriell K12 (NA12878).....	27
5.2.2.3.	Comparación de las tasas de error obtenidas para cada línea celular.....	28
5.2.2.4.	Genes más polimórficos	28
5.2.2.5.	Genes polimórficos relacionados con Neoplasias Mieloides Hereditarias, anemia de Fanconi y enfermedad de Blackfan-Diamond	29
5.2.3.	Zonas de mejora	32
6.	CONCLUSIONES.....	34
7.	BIBLIOGRAFÍA	35
8.	ANEXOS.....	39
8.1.	Anexo I.....	39
8.2.	Anexo II.....	48

Índice de Figuras

Figura 1. Funciones celulares de las proteínas codificadas por los genes de predisposición implicados en HMMSs (del inglés, <i>hereditary myeloid malignancy syndromes</i>). Reimpreso de: The University of Chicago Hematopoietic Malignancies Cancer Risk Team. <i>Blood</i> . 2016. (Churpek & Godley, 2016)	2
Figura 2. Regiones más frecuentemente mutadas de los genes de predisposición a Neoplasias Mieloides familiares: a CEBPA, b RUNX1, c GATA2, d DDX41, e ETV6 y f ANKRD26. Reimpreso de Király et al. <i>Pathol. Oncol. Res.</i> 2018 (Király et al., 2018).....	7
Figura 3. Ejemplo de algoritmo para la identificación de los pacientes apropiados para la clínica de neoplasias hematológicas hereditarias. AA indica anemia aplásica; LMA, leucemia mieloide aguda; SMD, síndrome mielodisplásico. Adaptado de DiNardo et al. <i>Cancer</i> . 2018. (DiNardo et al., 2018)	8
Figura 4. Hallazgos genómicos con potencial para las Neoplasias Mieloides con predisposición en línea germinal y las enfermedades hematológicas hereditarias. Reimpreso de DiNardo et al. <i>Cancer</i> . 2018. (DiNardo et al., 2018)	10
Figura 5. Esquema de trabajo del enriquecimiento diana para NGS <i>SureSelect^{QXT}</i> . Imagen adaptada de “ <i>SureSelect^{QXT} Automated Target Enrichment for Illumina Multiplexed Sequencing. Automated using Agilent NGS Bravo, 2015.</i> ” (AGILENT TECHNOLOGIES, 2015).....	15
Figura 6. Esquema resumen del protocolo llevado a cabo para los ensayos Qubit. (QUBIT 4 FLUOROMETER - ES, S.F.).....	16
Figura 7. Librería de precaptura preparada a partir de muestras de ADN _g analizado utilizando un ensayo <i>D1000 ScreenTape</i>	17
Figura 8. Análisis del ADN capturado amplificado empleando una <i>TapeStation 2100</i>	18

Índice de Tablas

Tabla 1. Clasificación de la OMS de Neoplasias Mieloides con predisposición de la línea germinal. (Steven H. Swerdlow et al., 2016)	3
Tabla 2. Rangos cuantitativos para cada uno de los ensayos Qubit. Adaptación de: (QUBIT 4 FLUOROMETER - ES, S.F.)	16
Tabla 3. Listado de genes incluidos en el panel que tienen pseudogenes conocidos. Se incluye información sobre su NM, el cromosoma donde se localizan y la información extraída de la base de datos Gene del NCBI, sobre la presencia de pseudogenes y la información acerca de los mismos para cada uno. Chr: <i>chromosome</i> , del inglés, cromosoma. NM: identificador de la secuencia de referencia.	20
Tabla 4. Listado de genes incluidos en el panel que poseen mutaciones intrónicas profundas que es necesario cubrir. Se indican el gen, la localización genómica hg19, variantes según la HGVS (<i>Human Genome Variation Society</i>), el identificador de la secuencia de referencia en RefSeq y el número RS. Chr: <i>chromosome</i> , del inglés, cromosoma. HGVS: <i>Human Genome Variation Society</i> . IVS: <i>intervening sequence</i> . ND: no disponible. NM: identificador de la secuencia de referencia. RS: referencia del SNP.....	22
Tabla 5. Listado de genes incluidos en el panel que poseen mutaciones en las regiones reguladoras. Se incluye el nombre del gen, su NM, su localización cromosómica y las regiones a cubrir en cada uno de ellos. Chr: <i>chromosome</i> , del inglés, cromosoma. NM: identificador de la secuencia de referencia. UTR: <i>untranslated region</i>	23
Tabla 6. Valores de cobertura obtenidos para cada uno de los ensayos llevados a cabo para cada línea celular, en base 1.....	24
Tabla 7. Resumen de la información obtenida tras el ensayo realizado con la línea celular Coriell D1, en que se indica para cada uno de los ensayos (M1, M2 y M3), el número de posiciones detectadas, los falsos positivos, el porcentaje de error y el porcentaje de verdaderos positivos, además de la media para todos los ensayos de los últimos dos.	25
Tabla 8. Tipo de variantes detectadas en la línea celular Coriell D1 en absoluto y porcentaje de cada tipo sobre el total.....	26
Tabla 9. Tasa de error por tipo de variante para cada uno de los 3 ensayos realizados con la línea celular Coriell D1.	26
Tabla 10. Resumen de la información obtenida tras el ensayo realizado con la línea celular Coriell K12, en que se indica para cada uno de los ensayos (M1, M2 y M3), el número de posiciones detectadas, los falsos positivos, el porcentaje de error y el porcentaje de verdaderos positivos, además de la media para todos los ensayos de los últimos dos.	27
Tabla 11. Tipo de polimorfismo de las posiciones detectadas en la línea celular Coriell K12 en absoluto y porcentaje de cada tipo sobre el total.	27
Tabla 12. Tasa de error por tipo de variante para cada uno de los 3 ensayos realizados con la línea celular Coriell K12.	28

Tabla 13. Genes más polimórficos para las líneas celulares Coriell D1 y K12. Se destacan aquellos genes que coinciden en ambas líneas celulares.....	29
Tabla 14. Genes polimórficos asociados a las NMH y número de variantes detectadas en cada uno para las líneas celulares Coriell D1 y Coriell K12.	30
Tabla 15. Genes polimórficos asociados a la anemia de Fanconi y número de variantes detectadas en cada uno para las líneas celulares Coriell D1 y Coriell K12.....	31
Tabla 16. Genes polimórficos asociados a la anemia de Fanconi y número de variantes detectadas en cada uno para las líneas celulares Coriell D1 y Coriell K12.....	32

Anexos

Tabla Anexo I. Diseño del panel de genes dirigido a neoplasias mieloides hereditarias: genes incluidos en el panel e información sobre su localización, la condición patológica asociada y su NM. Información extraída de las bases de datos Gene, ClinVar y PubMed, del NCBI. Chr: <i>chromosome</i> , del inglés, cromosoma. ND: no disponible. NM: identificador de la secuencia de referencia.	39
Tabla Anexo II. Zonas de mejora del panel de genes dirigido. Se incluyen la localización cromosómica, las posiciones concretas iniciales y finales, el gen correspondiente y el NM junto con el exón correspondiente. Chr: <i>chromosome</i> , del inglés, cromosoma. NM: identificador de la secuencia de referencia.	48

Lista de abreviaturas

- AA:** anemia aplásica
- ACMG:** *American College of Medical Genetics*
- ADN:** Ácido desoxirribonucleico
- ADNg:** Ácido desoxirribonucleico genómico
- ARN:** Ácido ribonucleico
- BMF:** insuficiencia de la médula ósea
- BR:** *Broad Range*
- Chr:** *chromosome*
- HMMS:** *Hereditary Myeloid Malignancy Syndrome*
- HS:** *High Sensitivity*
- INDELS:** inserciones y deleciones
- IVS:** *Intervening sequence*
- Kb:** kilobases
- LLA:** Leucemia linfoblástica aguda
- LLC:** leucemia linfocítica crónica
- LMA:** Leucemia mieloide aguda
- LMC:** leucemia mieloide crónica
- MAF:** frecuencia de alelo menor
- ND:** no disponible
- NGS:** *Next-generation Sequencing*
- NM:** identificador de la secuencia de referencia
- NMH:** Neoplasias Mieloides Hereditarias
- NMP:** neoplasias mieloproliferativas familiares
- OMS:** Organización Mundial de la Salud
- Pb:** pares de bases
- PCR:** *Polimerase Chain Reaction*
- RS:** referencia del SNP
- S.F.:** sin fecha
- SMD:** Síndrome Mielodisplásico
- SNP:** *Single Nucleotide Polimorfism*
- UTR:** *untranslated region*

1. INTRODUCCIÓN

Las Neoplasias Mieloides tratan de un conjunto de enfermedades donde las células sanguíneas producidas en la médula ósea no son capaces de diferenciarse, perpetuándose en su estado indiferenciado. De este modo, los glóbulos rojos, blancos y plaquetas no son capaces de evolucionar hacia células sanguíneas maduras. Este TFG se centra en las Neoplasias Mieloides, que incluyen la leucemia mieloide aguda (LMA) y el síndrome mielodisplásico (SMD).

La LMA es el tipo más común de leucemia *de novo* en adultos. Es causada por células mieloblásticas malignas con mutaciones adquiridas que se expanden clonalmente y evitan la diferenciación después (West et al., 2014). El SMD es una neoplasia maligna de médula ósea donde una hematopoyesis alterada produce células precursoras con displasia morfológica y citopenias de células de la sangre periférica. (Vardiman et al., 2009) El SMD generalmente tiene un inicio tardío y, a veces, evoluciona a LMA (West et al., 2014). La LMA y los SMD de carácter familiar han sido considerados clásicamente como neoplasias muy raras, vinculadas principalmente a trastornos pediátricos. Sin embargo, los recientes avances tecnológicos ligados al desarrollo de la NGS han permitido identificar mutaciones germinales con una frecuencia mayor de lo esperada en pacientes con estas neoplasias de estirpe mieloide. De hecho, estas entidades han sido incluidas como categorías diferenciadas en la reciente Clasificación de la Organización Mundial de la Salud 2016 (Steven H. Swerdlow et al., 2016). Por todo ello, es necesario desarrollar una estrategia de NGS que permita recoger la máxima información de la línea germinal de estos pacientes y compararla con la de la muestra de médula ósea pareada en el momento del diagnóstico, así como con la historia clínica familiar.

El diagnóstico clínico de las HMMSs (*Hereditary Myeloid Malignancy Syndromes*) es complicado, principalmente en la edad adulta, en parte, como consecuencia de la asociación que presentan muchos de estos genes con los casos de neoplasias mieloides adquiridas. Para ello, es imprescindible realizar por parte del facultativo una anamnesis completa que recoja antecedentes de hemopatías, tales como trombocitopenias u otras citopenias, cáncer, desórdenes de coagulación, o enfermedades pulmonares o hepáticas en los familiares del paciente. El reconocimiento de un caso de HMMSs en la práctica clínica es imprescindible, ya que varía las guías de tratamiento y de seguimiento de estas familias, una mejor selección de donantes de médula ósea y consejo y asesoramiento genético al paciente y a la familia afectada. (Steven H. Swerdlow et al., 2016)

1.1. Neoplasias Mieloides Hereditarias

Dentro de las Neoplasias Mieloides Hereditarias (NMH), los trastornos más establecidos son los que presentan manifestaciones fenotípicas adicionales que a menudo se presentan en la infancia, como es el caso de los síndromes de insuficiencia de la médula ósea heredados, tales como la Anemia de Fanconi, el Síndrome de Shwachman-Diamond o la Disqueratosis Congénita. (García et al., 2018)

mecanismos moleculares y celulares. Y a medida que aumentan los proyectos de secuenciación, se van encontrando cada vez más y más genes relacionados.

Por todo ello, se terminó estableciendo la nueva categoría de “Neoplasias Mieloides con predisposición en línea germinal” en la revisión de 2016 de la Clasificación de Neoplasias Mieloides y leucemia aguda de la Organización Mundial de la Salud (OMS). (Steven H. Swerdlow et al., 2016)

1.1.1. Clasificación de la OMS de las Neoplasias Mieloides Hereditarias

La Organización Mundial de la Salud, en su revisión de 2016, ha decidido presentar una clasificación de las NMH en tres grandes grupos, donde encontramos Neoplasias Mieloides con predisposición en línea germinal sin desorden preexistente o disfunción de órganos, las que sí presentan desórdenes plaquetarios preexistentes y las que presentan otras disfunciones de órganos.

Además, dentro de cada categoría, encontramos diversos subgrupos, organizados tanto por genes mutados como por patología asociada. (Tabla 1)

Tabla 1. Clasificación de la OMS de Neoplasias Mieloides con predisposición de la línea germinal. (Steven H. Swerdlow et al., 2016)

Neoplasias Mieloides con predisposición en línea germinal
Neoplasias Mieloides con mutaciones predisponentes en línea germinal sin desorden preexistente o disfunción de órganos
LMA con mutación en <i>CEBPA</i> en línea germinal
Neoplasias Mieloides con mutación germinal en <i>DDX41</i> *
Neoplasias Mieloides con mutaciones predisponentes en línea germinal y desórdenes plaquetarios preexistentes
Neoplasias Mieloides con mutación germinal en <i>RUNX1</i> *
Neoplasias Mieloides con mutación germinal en <i>ANKRD26</i> *
Neoplasias Mieloides con mutación germinal en <i>ETV6</i> *
Neoplasias Mieloides con mutaciones predisponentes en línea germinal y disfunciones de órganos
Neoplasias Mieloides con mutación germinal en <i>GATA2</i>
Neoplasias Mieloides asociadas con el síndrome de fallo BM
Neoplasias Mieloides asociadas con desórdenes en la biología de los telómeros
Neoplasias Mieloides asociadas con Neurofibromatosis, Síndrome de Noonan o desórdenes parecidos al síndrome de Noonan
Neoplasias Mieloides asociadas con síndrome de Down*

*También se han notificado neoplasias linfoides con estas mutaciones de la línea germinal.

1.1.1.1. Neoplasias Mieloides con mutaciones predisponentes en línea germinal sin desórdenes plaquetarios o disfunción de órganos

1.1.1.1.1. LMA con mutación en *CEBPA* en línea germinal

El factor de transcripción *CEBPA* (del inglés, *CCAAT/enhancer binding protein-a*) está codificado por un gen que se sitúa en el cromosoma 19q13.1. Dicho factor de transcripción, que tiene un importante papel en el control del crecimiento celular y en la diferenciación granulocítica, presenta dos zonas diferenciadas donde se agrupan las mutaciones, que son las zonas N-terminal y C-terminal. (Nerlov, 2007) En el caso de la región C-terminal, se ha encontrado relación con mutaciones somáticas. Por contra, la región N-terminal agrupa las mutaciones de carácter germinal, y se trata en su mayoría de mutaciones del marco.

A pesar de ello, se ha relacionado la presencia de mutaciones de *CEBPA* en línea somática con la presencia de mutaciones en línea germinal entre el 5 y el 10 % de los casos de LMA. (Taskesen et al., 2011) La identificación de las mutaciones en línea germinal resulta de gran importancia a la hora de seleccionar potenciales donantes de células hematopoyéticas, para tener en cuenta que donante y receptor no sean ambos portadores de la misma mutación. (Xiao et al., 2010)

Además, se ha demostrado que la edad media de desarrollo de la LMA es a los 24'5 años, con posibilidad de recaídas, pero se concluyó que este tipo de leucemia presenta un pronóstico favorable. (Taskesen et al., 2011)

1.1.1.1.2. Neoplasias Mieloides con mutación germinal en *DDX41*

El cromosoma 5q35 alberga el gen que codifica para la helicasa *DDX41* (del inglés, *DEAD/H-box helicase*). Esta familia de proteínas tiene función en la hematopoyesis, pues se expresa en las células CD14+, CD33+ y CD34+. Además, las mutaciones germinales en la región N-terminal con desplazamientos del marco de lectura y deleciones en el supresor tumoral *DDX41* que conllevan la pérdida de su función o que favorecen el crecimiento celular de las Neoplasias Mieloides, son las que se presentan con mayor frecuencia. (Polprasert et al., 2015)

Además, las mutaciones en *DDX41* presentan un patrón de herencia autosómico dominante y se asocian a diversas Neoplasias Mieloides y linfoides. Un estudio demostró que los portadores de dichas mutaciones portaban también mutaciones en línea germinal en la mitad de los casos. (Lewinsohn et al., 2016)

Se ha visto que la edad de aparición de estas enfermedades es tardía y detectable una vez se ha desarrollado el proceso maligno, pues no presentan ningún rasgo hematológico diferenciable para su temprano diagnóstico, por lo que para ello sería necesario presentar un alto índice de sospecha y estudiar la historia familiar exhaustivamente. (Bannon & Dinardo, 2016)

1.1.1.1.3. Neoplasias Mieloides con mutación germinal del gen *SRP72**

SRP72 (signal recognition particle 72) codifica la subunidad de 72 kDa de la partícula de reconocimiento de señal (SRP), un complejo ribonucleoproteínico responsable de detener la traducción de proteínas secretoras o extracelulares y dirigirlas al retículo endoplásmico. (Kirwan et al., 2012)

Las mutaciones germinales en *SRP72* se asocian a aplasia y predisposición al desarrollo de formas hereditarias de SMD de herencia dominante. Los primeros casos descritos fueron en dos familias no relacionadas en el año 2011; individuos de ambas familias presentaban desórdenes hematológicos además de pérdidas de audición o anomalías audiovestibulares (Godley, 2014; Babushok et al., 2016)

Debido a la escasez de casos descritos, se desconoce la incidencia, el riesgo de desarrollar una neoplasia o qué indicaciones clínicas seguir en estas familias.

1.1.1.2. *Neoplasias Mieloides con mutaciones predisponentes de la línea germinal y desórdenes plaquetarios preexistentes*

1.1.1.2.1. Neoplasias Mieloides con mutación germinal en *RUNX1*

El gen *RUNX1* se localiza en el cromosoma 21q22. Se caracteriza por su importante papel en el control de la hematopoyesis, función que realiza a través de la regulación de otros genes, entre los que se incluyen moléculas de señalización y factores de crecimiento. La mutación de este gen provoca defectos tanto cuantitativos como cualitativos en las plaquetas, además de una pobre expresión del receptor MPL. (Lordier et al., 2012; Jalagadugula et al., 2010)

Esto es debido a que la maduración megacariopoyética se ve afectada con las mutaciones de este gen, entre las que podemos encontrar mutaciones puntuales, sin sentido, duplicaciones intragénicas, translocaciones y desplazamientos de pautas de lectura. (Schnittger et al., 2011) (Sakurai et al., 2016; Antony-Debré et al., 2016) Pero para el desarrollo de la enfermedad, es necesaria la presencia de otras mutaciones colaboradoras. (Hayashi et al., 2017)

En cuanto al patrón de herencia, es autosómico dominante y presenta ciertas manifestaciones hematológicas asociadas, como son la presencia de trombocitopenia de ligera a moderada, sangrados mucosos y una alta predisposición a padecer tanto LMA como SMD. (Nishimoto et al., 2010) También se pueden observar manifestaciones no hematológicas, entre las que destaca el desarrollo de eccemas.

En cuanto a la morfología plaquetaria asociada, la lámina periférica aparece de tamaño normal, lo que se asocia a cambios dismegacariopoyéticos en la médula ósea. Además, la función plaquetaria se ve afectada en las pruebas de agregación. (Liew & Owen, 2011; Ok et al., 2016)

1.1.1.2.2. Neoplasias Mieloides con mutación germinal en *ANKRD26*

Las mutaciones del gen *ANKRD26* (del inglés, *ankirin repeat domain*) en su región 5' no traducida, se localizan en el cromosoma 10p12. Dichas mutaciones son las causantes del desarrollo de trombocitopenia tipo 2, síndrome que presenta un patrón de herencia autosómico dominante. (Bannon & Dinardo, 2016)

Diversos estudios han concluido que, efectivamente, las mutaciones del gen *ANKRD26* se asocian con trombocitopenias y hemorragias generalmente ligeras. Además, el tamaño de la lámina periférica de las plaquetas era normal, ligeramente pálidas, y presentaban frecuentemente dismegacariopoyesis en la médula ósea, función plaquetaria normal, elevados

niveles de trombopoyetina y una alta predisposición al desarrollo de LMA, SMD, y, en menor medida, LLC (leucemia linfocítica crónica) y LMC (leucemia mieloide crónica). (Bannon & Dinardo, 2016; Nishimoto et al., 2010)

1.1.1.2.3. Neoplasias Mieloides con mutación germinal en *ETV6*

El gen *ETV6* se localiza en el brazo corto del cromosoma 12, y codifica un factor de represión de la transcripción. Además, su patrón de herencia es autosómico dominante. Las mutaciones en *ETV6* se han relacionado con el desarrollo de trombocitopenia tipo 5, sangrados mucosos, macrocitosis, y también predisposición a desarrollar LLA, LMA y SMD, además de otros cánceres como de piel y de colon. (Noetzli et al., 2015; Zhang et al., 2015)

1.1.1.3. *Neoplasias Mieloides con mutaciones predisponentes de la línea germinal y disfunción de órganos*

1.1.1.3.1. Neoplasias Mieloides con mutación germinal en *GATA2*

El gen *GATA2* se localiza en el cromosoma 3q21, y forma parte de la familia de factores de transcripción *GATA*. Se han asociado diversos tipos de mutaciones al gen *GATA2*, como son inserciones, pequeñas y grandes deleciones, mutaciones puntuales, mutaciones de cambio de pauta de lectura y mutaciones de terminación prematura. (Hahn et al., 2011; Pasquet et al., 2013)

Las mutaciones en línea germinal de dicho gen aumentan el riesgo de padecer SMD y LMA. Además, se asocian también al desarrollo de desórdenes complejos, como el MonoMAC (caracterizado por una ausencia total de monocitos), entre otros. (Hsu et al., 2012; M.A. et al., 2014)

También cabe destacar que un estudio ha concluido que las mutaciones en el gen *GATA2* en línea germinal constituyen el factor que más predispone al desarrollo de SMD en población pediátrica. (Wlodarski et al., 2016) Aunque tal y como refleja la Figura 2, hay más genes que poseen regiones cuyas mutaciones más frecuentes predisponen a NMH.

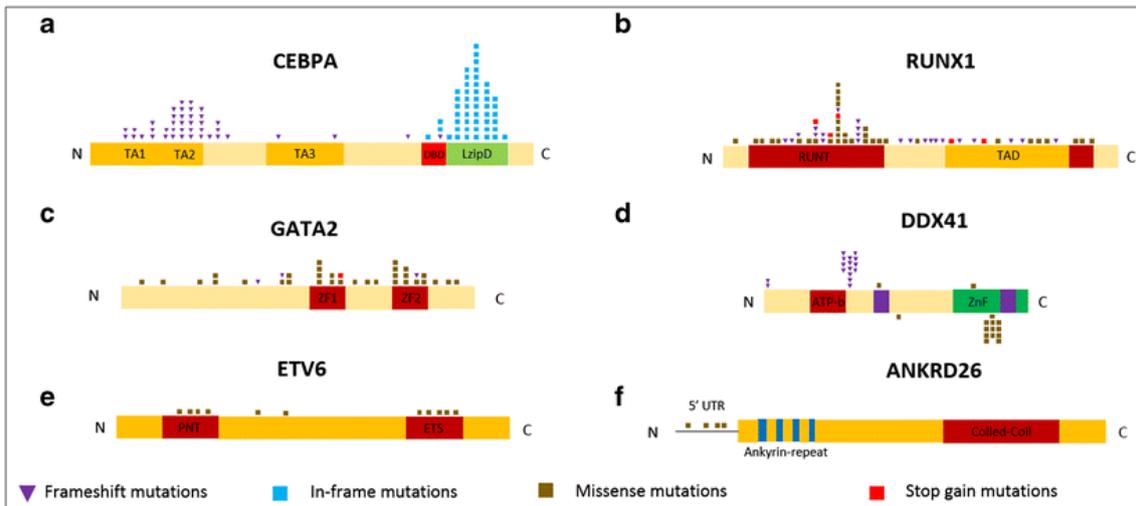


Figura 2. Regiones más frecuentemente mutadas de los genes de predisposición a Neoplasias Mieloides familiares: a CEBPA, b RUNX1, c GATA2, d DDX41, e ETV6 y f ANKRD26. Reimpreso de Király et al. Pathol. Oncol. Res. 2018 (Király et al., 2018)

1.1.2. Diagnóstico y manejo de pacientes con Neoplasias Mieloides Hereditarias

El diagnóstico de este tipo de enfermedades puede resultar complejo, pues muchos de los genes involucrados participan no solo en el desarrollo de NMH, sino también en las Neoplasias Mieloides adquiridas. Por tanto, para un correcto diagnóstico, sería necesario evaluar detalladamente los historiales médicos tanto individuales como familiares, tratando de buscar antecedentes que puedan hacer sospechar de la presencia de las neoplasias, como otras neoplasias hematológicas u otros tipos de cáncer. Además, también se deben tener en cuenta otros síntomas no cancerosos, que tal y como hemos comentado anteriormente, se relacionan con la presencia de ciertas mutaciones en línea germinal, como son las citopenias graves, trombocitopenias, episodios de sangrado y disfunciones plaquetarias.

A pesar de esto, el diagnóstico de las NMH no debe basarse únicamente en precedentes, puesto que, en muchos casos, los historiales familiares no están disponibles. En la Figura 3, podemos ver un ejemplo de algoritmo para la identificación de los pacientes apropiados para la clínica de neoplasias hematológicas hereditarias.

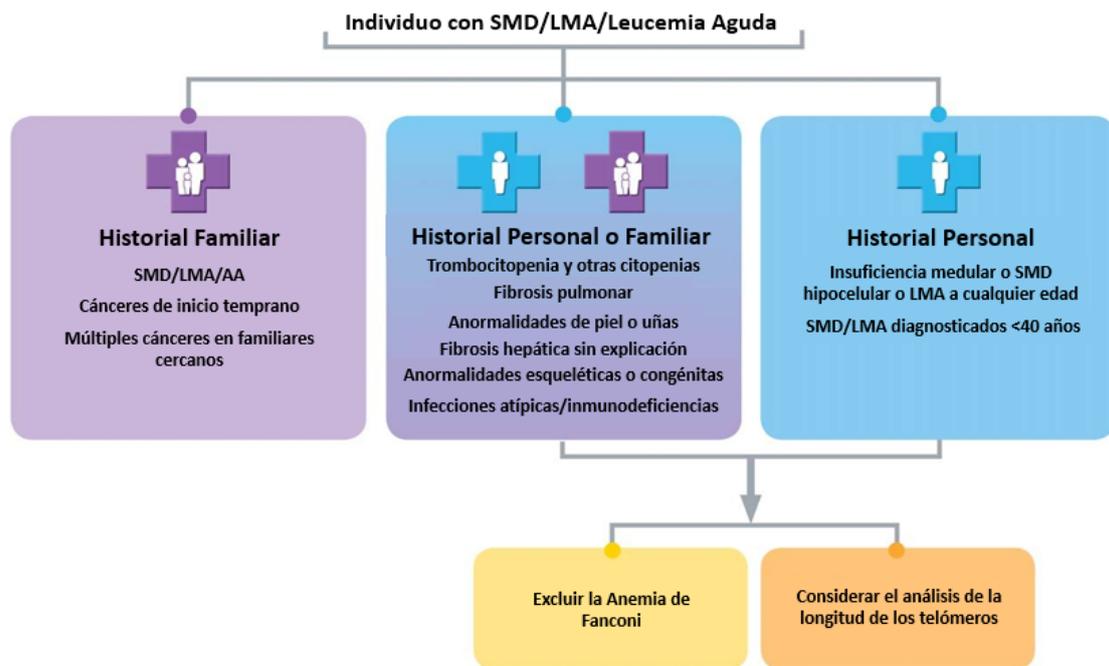


Figura 3. Ejemplo de algoritmo para la identificación de los pacientes apropiados para la clínica de neoplasias hematológicas hereditarias. AA indica anemia aplásica; LMA, leucemia mieloide aguda; SMD, síndrome mielodisplásico. Adaptado de DiNardo et al. *Cancer*. 2018. (DiNardo et al., 2018)

Es por ello por lo que se han implementado los paneles de genes NGS comerciales y *custom* con el objetivo de diagnosticar LMA y SMD, pero estos suelen incluir genes con mutaciones somáticas, por lo que se hace evidente la necesidad de diseñar y poner a punto paneles de genes que contengan aquellos genes cuyas mutaciones en línea germinal predispongan al desarrollo de las NMH.

En conclusión, el diagnóstico de las NMH debe basarse tanto en estudios del historial individual y familiar, como en el estudio morfológico y citogenético de muestras de sangre periférica y médula ósea, realizando un análisis molecular mediante un panel de genes dirigido que incluya los genes con predisposición para la detección de mutaciones en línea germinal, donde destacan los genes *CEBPA*, *DDX41*, *RUNX1*, *ANKRD26*, *ETV6*, *GATA2* y *SRP72*.

En cuanto al manejo de los pacientes, cabe destacar la importancia de llevar a cabo una selección óptima del donante, a partir de una exhaustiva evaluación para descartar la presencia de mutaciones en línea germinal. Los familiares cercanos pueden ser portadores, y es necesario minimizar el riesgo de elegir un donante afectado, por lo que se suele preferir la elección de donantes no relacionados.

Además, el diagnóstico temprano de las NMH es crucial para la elección del tratamiento y supervisión del paciente. Tanto los pacientes como sus familiares se incluyen en programas de vigilancia y son informados sobre su situación y la necesidad de llevar a cabo un control continuo.

1.2. Uso de paneles de genes dirigidos para detectar mutaciones

La utilización de paneles de genes dirigidos es algo común en la práctica clínica, ya que a menudo se encuentran múltiples mutaciones con diversas implicaciones pronósticas y terapéuticas en el diagnóstico inicial de pacientes que padecen Neoplasias Mieloides. Por ello, resultan ser métodos con una alta rentabilidad diagnóstica. Además, al detectar simultáneamente mutaciones en genes diferentes, constituye un enfoque más completo a la hora de tomar decisiones sobre el tratamiento adecuado que empleando las técnicas tradicionales de PCR y secuenciación. (DiNardo et al., 2018)

Los genes incluidos en los paneles de genes dirigidos a la detección de tumores hematológicos malignos difieren según la plataforma e institución, pero la mayoría suelen compartir alrededor de 25 genes que son los que se encuentran mutados en línea germinal del paciente con mayor frecuencia en las leucemias agudas y en las Neoplasias Mieloides (*CEBPA*, *GATA2*, *RUNX1*, *TP53*, entre otros). Por tanto, el diseño de los paneles de genes dirigidos como una herramienta de diagnóstico, ofrece la oportunidad de determinar el origen e importancia clínica de mutaciones en línea germinal en paneles diseñados principalmente para la detección y análisis de mutaciones somáticas. (DiNardo et al., 2018)

Sin embargo, las implicaciones de una mutación pueden diferir según la edad, sexo, antecedentes de los pacientes u otras características, lo que dificulta la evaluación del riesgo de la presencia de estas. Además, a la hora de distinguir entre mutaciones germinales y adquiridas, se necesita una evaluación del tejido de línea germinal con el fin de ofrecer un diagnóstico de predisposición al desarrollo de cáncer.

También cabe destacar que, una vez obtenida la lista de genes afectados tras la NGS, es necesario evaluar su importancia clínica a partir de un posterior análisis bioinformático, con la consulta de diversas bases de datos, muchas de ellas públicas.

En resumen, la NGS dirigida puede ser útil para guiar la selección del tratamiento, proporcionar información sobre el pronóstico y la evolución del tumor, evitar la resistencia al tratamiento, promover el desarrollo de nuevos fármacos y comprender completamente los mecanismos moleculares subyacentes a la progresión del proceso neoplásico. En la Figura 4 podemos ver un esquema resumen de hallazgos genómicos con potencial para las Neoplasias Mieloides con predisposición en línea germinal y las enfermedades hematológicas hereditarias.

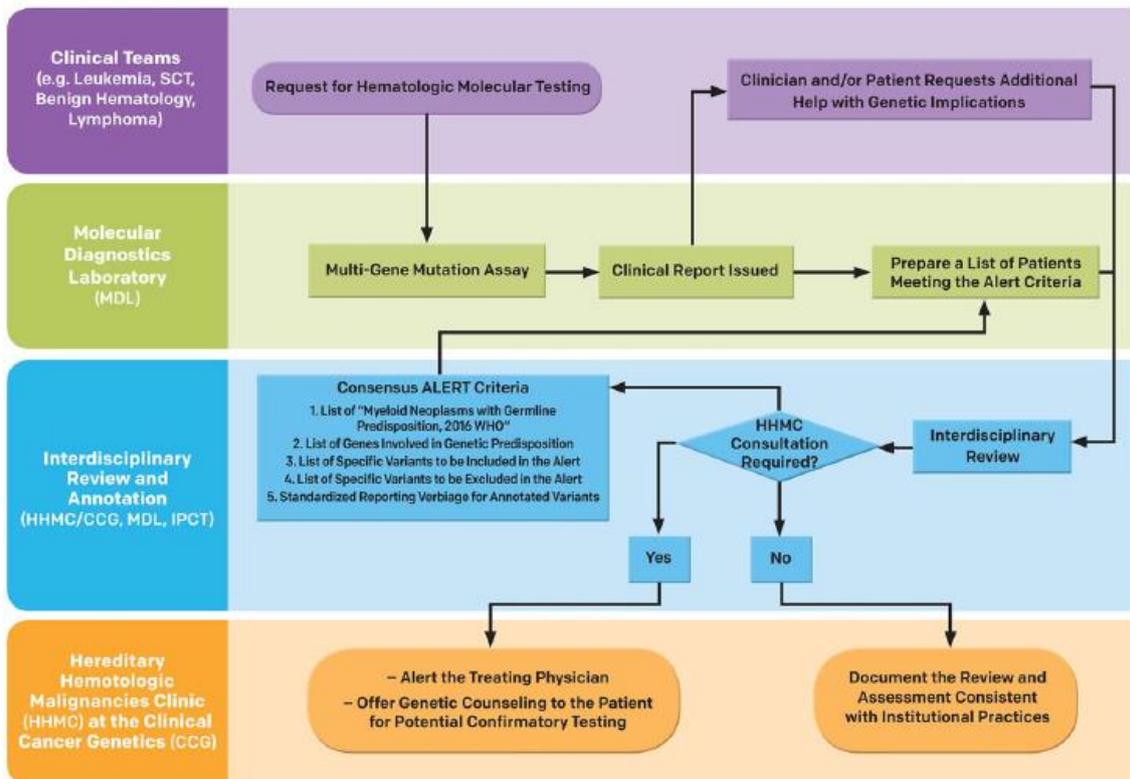


Figura 4. Hallazgos genómicos con potencial para las Neoplasias Mieloides con predisposición en línea germinal y las enfermedades hematológicas hereditarias. Reimpreso de DiNardo et al. *Cancer*. 2018. (DiNardo et al., 2018)

2. HIPÓTESIS

Debido al aumento de la identificación de mutaciones genéticas en línea germinal que predisponen al desarrollo de neoplasias mieloides, se hace necesaria la implementación de técnicas que permitan detectar este tipo de variantes, con el objetivo de diagnosticar la enfermedad y ofrecer un tratamiento lo más efectivo posible al paciente.

Por ello, se propone el diseño y validación de un panel de genes *custom* dirigido a NMH, una herramienta de diagnóstico que hace uso de las avanzadas tecnologías de NGS con el objetivo de proporcionar una técnica de monitoreo molecular que logre cubrir las actuales necesidades clínicas.

3. OBJETIVOS

En el presente proyecto se pretende rediseñar y validar un panel de genes que permita abordar simultáneamente el estudio global de un gran número de alteraciones de forma masiva en estos pacientes e identificar nuevos casos de LMA/SMD familiar. Los objetivos específicos de esta propuesta de investigación son:

1. Búsqueda bibliográfica exhaustiva y actualizada sobre el estudio a nivel genético de las NMH.
2. Rediseño de un panel de genes por NGS para el estudio de mutaciones germinales en todos aquellos genes asociados a HMMSs en pacientes adultos con sospecha de LMA o SMD familiar en el momento del diagnóstico.
3. Validación del panel utilizando líneas celulares Coriell como patrones para NGS, obteniendo los correspondientes valores de sensibilidad y especificidad del panel diseñado.
4. Implantar en el laboratorio clínico el panel de genes por NGS para diagnosticar pacientes con LMA/SMD familiar.

4. MATERIALES Y MÉTODOS

4.1. Diseño del panel de genes dirigido

4.1.1. Búsqueda bibliográfica

Para determinar qué genes incluir en el panel, se partió de la información de que se disponía sobre la versión anterior del panel diseñado en el Instituto de Investigación Sanitaria La Fe, descartando y añadiendo genes tras contrastar su utilidad para el estudio. Además, se realizó una búsqueda bibliográfica exhaustiva, seleccionando aquellos artículos más recientes que contenían información al respecto (por ejemplo DiNardo et al., 2018; Desai et al., 2017; Tawana et al., 2018; Drazer et al., 2018; Akpan et al., 2018).

4.1.2. Recopilación de información de bases de datos

Para cada gen, se han utilizado las bases de datos OMIM, ClinVar, PubMed y ENSEMBL para anotar la condición patológica asociada a cada uno, su localización en el cromosoma y su identificador de la secuencia de referencia (NM). Además, para cada uno de ellos, se ha analizado la presencia de pseudogenes, aumentándose la densidad de sondas de esa región en aquellos genes en los que se confirmaba. Este hecho es muy relevante, ya que, si una sonda mapea en dos regiones (exón y pseudoexon), se perdería cobertura y la profundidad real de lectura disminuiría considerablemente.

Con el fin último de definir convenientemente las regiones génicas a incluir en el panel, se han anotado cada una de las variantes descritas previamente en las bases de datos HGMD y ClinVar, teniendo especial interés aquellas que se localizaban en regiones de splicing o en regiones reguladoras de la expresión del gen.

4.1.3. Diseño final del panel de genes

El panel de genes se ha diseñado utilizando el sistema *SureDesign Custom Design Tool* (Agilent Technologies Inc.). Este diseño incluye un total de 177 genes relacionados con NMH, incluyendo los síndromes familiares de predisposición a SMD/LMA (entre otros: *RUNX1*, *ETV6*, *CEBPA*, *GATA2*, *DDX41*, *ANKRD26*, *SAMD9*, *SAMD9L*) y los síndromes hereditarios con insuficiencia de médula ósea con predisposición a SMD/LMA, como anemia de Fanconi (*BRCA2*, *FANCF*, *FANCG*), disqueratosis congénita (*CTC1*, *DKC1*), anemia de Diamond-Blackfan (*RPS19*, *RPL5*), Síndrome de Schwachman-Diamond (*SBD*), anemia sideroblástica congénita (*STEAP3*, *HSPA9*), neutropenia congénita grave (*HAX1*, *GFI1*), entre otros. Este panel de genes permitirá capturar las regiones codificantes completas y las regiones flanqueantes de los genes seleccionados, y será planteado para detectar >99% de las mutaciones.

Se incluye junto al trabajo el ANEXO I, donde aparecen todos los genes empleados para el diseño del panel junto con la información recopilada de cada uno de ellos en la Tabla Anexo I. Por otro lado, en el subapartado 5.1. del apartado de Resultados y discusión, se muestran las tablas obtenidas que incluyen los genes que presentan especificaciones especiales (presencia de pseudogenes, regiones intrónicas profundas de interés y genes con mutaciones patológicas en las zonas reguladoras).

4.2. Validación del panel de genes dirigido

Para la validación del panel de genes diseñado, se siguieron las directrices de las guías de buenas prácticas actuales en la implementación de la secuenciación masiva a nivel clínico, el *EuroGentest Group* (EUROGENTEST, S.F.) y el *American College of Medical Genetics* (ACMG) (AMERICAN COLLEGE OF MEDICAL GENETICS AND GENOMICS (ACMG), S.F.).

Los ensayos de NGS necesitan ser validados con el objetivo de cumplir una serie de especificaciones y características que permitan comparaciones entre laboratorios. Tras el diseño informático del panel, es necesario validar la primera versión del mismo con el objetivo de utilizarlo para testar su funcionamiento y la veracidad de los resultados obtenidos, comparándolos con los teóricamente esperados.

La ACMG recomienda el uso de muestras de referencia a los laboratorios que utilicen NGS que estén previamente secuenciadas, normalizadas y de alta calidad, para tener una base de datos sobre la que comparar los resultados obtenidos durante la validación. Con este fin, se decidió emplear las líneas celulares Coriell D1 (NA12144) y Coriell K12 (NA12878), proporcionadas por el Instituto de Investigación Sanitaria La Fe, para la validación analítica de secuenciación NGS y del análisis bioinformático, en la detección de SNPs (*Single Nucleotide Polimorfisms*) e INDELS (pequeñas inserciones/deleciones <20nt). Estas líneas celulares son idóneas porque está establecido un set de mutaciones localizadas y almacenadas en bases de datos en las muestras de referencia utilizadas. Con el objetivo de minimizar la tasa de error, cada línea celular se analizó por triplicado (M1, M2 y M3). Es necesario destacar que cada triplicado se ha procesado de manera independiente, como un ensayo ciego, utilizándose las muestras de referencia.

La utilización de líneas celulares previamente estudiadas en la validación del panel resulta útil debido a que se trabaja con ambas líneas detectando qué mutaciones son reconocidas por el panel y cuáles no, comparando con los datos registrados. De este modo, se logra calcular la tasa de error del panel, obteniendo una estimación de la sensibilidad del mismo. Además, se obtiene una aproximación de la cobertura real alcanzada con el panel, y con ello, se estiman las zonas mal cubiertas. Esto es interesante debido a que la información obtenida del estudio de las zonas no cubiertas se utilizará para tenerlo en cuenta en el diseño de la próxima versión del panel, que incluirá dichas posiciones cromosómicas como zonas de mejora.

4.2.1. Preparación del ADN de las líneas celulares

El ADN de las líneas celulares liofilizado se reconstituyó con tampón TAE 1x (tampón Tris-acetato EDTA) a una concentración recomendada de 200 ng/μl. Se realizaron alícuotas, en nuestro caso de 20 μl, y se conservaron para su posterior utilización en un congelador de -20°C.

4.2.2. Preparación de la librería enriquecida

La preparación de la librería y el enriquecimiento diana se realizó a través de “*SureSelect^{QXT} Automated Target Enrichment for Illumina Multiplexed Sequencing, 2015.*” de acuerdo con las instrucciones del fabricante. Todo el procedimiento se resume en la Figura 5.



Figura 5. Esquema de trabajo del enriquecimiento diana para NGS *SureSelect^{QXT}*. Imagen adaptada de “*SureSelect^{QXT} Automated Target Enrichment for Illumina Multiplexed Sequencing. Automated using Agilent NGS Bravo, 2015.*” (AGILENT TECHNOLOGIES, 2015)

4.2.2.1. Preparación de la librería de DNA genómico

La preparación de la librería de ADN genómico (ADNg) requiere de un paso crítico de cuantificación y normalización a una concentración final de 10 ng/ μ l, lo que asegura una fragmentación del ADN óptima. La cuantificación se realizó mediante dos ensayos fluorimétricos en serie: el ensayo *Qubit dsDNA BR* y el ensayo *Qubit dsDNA HS (ThermoFisher Scientific)*. En la Figura 6, se muestra un esquema resumen del protocolo llevado a cabo para los ensayos Qubit, y en la Tabla 2, los rangos cuantitativos para cada uno de los ensayos Qubit.

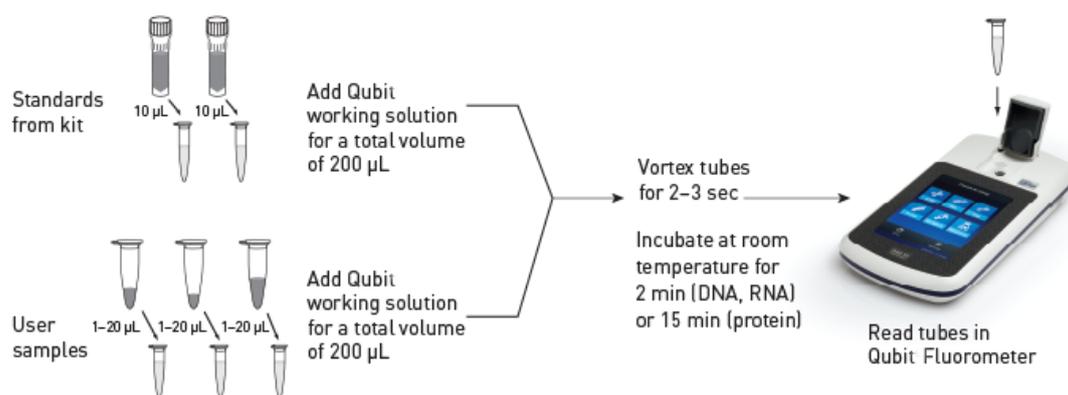


Figura 6. Esquema resumen del protocolo llevado a cabo para los ensayos Qubit. (QUBIT 4 FLUOROMETER - ES, S.F.)

Tabla 2. Rangos cuantitativos para cada uno de los ensayos Qubit. Adaptación de: (QUBIT 4 FLUOROMETER - ES, S.F.)

Producto	Concentración inicial de muestra	Rango de cuantificación
Ensayos de Amplio rango (<i>Broad Range</i>)		
Kit de ensayo Qubit <i>dsDNA BR</i>	100 pg/µl – 1000 ng/µl	2 – 1000 ng
Ensayos de alta sensibilidad (<i>High Sensitivity</i>)		
Kit de ensayo Qubit 1x <i>dsDNA HS</i>	10 pg/µl – 100 ng/µl	0'2 – 100 ng

Para llevar a cabo el proceso, se emplearon 5 ng de ADN que se fragmentaron enzimáticamente con una enzima transposasa al mismo tiempo que se añadieron adaptadores a los extremos de los fragmentos. Las muestras de ADN con los adaptadores marcados se purificaron utilizando perlas *AMPure XP*. Estas muestras de ADN ligadas con adaptador se amplificaron mediante PCR utilizando la polimerasa de ADN de fusión *Herculase II*, y se purificaron adicionalmente utilizando perlas *AMPure XP*.

La cantidad y calidad de la biblioteca de ADN se evaluó utilizando la *Agilent 4200 TapeStation* y una *D1000 ScreenTape (Agilent Technologies Inc.)* para determinar con precisión el tamaño del fragmento de ADN, que se muestra en la Figura 7 (300-350 pb).

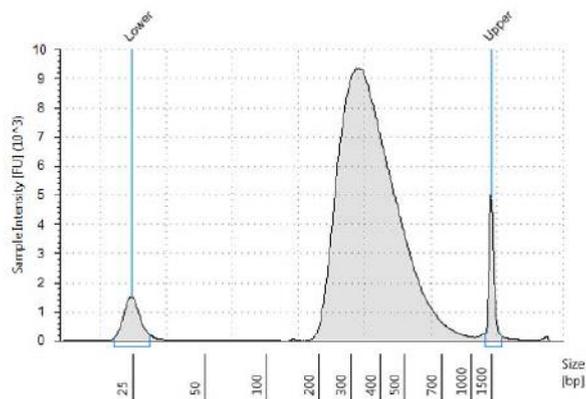


Figura 7. Librería de precaptura preparada a partir de muestras de ADN_g analizado utilizando un ensayo *D1000 ScreenTape*.

4.2.2.2. Hibridación

Tras la preparación de la librería de ADN_g, el siguiente paso consistió en la hibridación de esta con la librería de captura, con el objetivo de enriquecer las regiones específicas del genoma. Para ello, en primer lugar, se normalizaron las librerías de ADN_g ligadas con adaptador a 750 ng. Y después se hibridaron la librería de ADN_g marcada con el adaptador y la librería de captura. Esta última está compuesta por sondas de ARN marcadas con biotina. Este tipo de sondas de ARN se utiliza porque genera una mayor estabilidad de unión ADN-ARN que la que se daría con uniones ADN-ADN.

El siguiente paso consistió en la captura del híbrido, utilizando perlas Dyna-Beads (*Dynal MyOne Streptavidin T1 magnetic beads*; son perlas magnéticas cubiertas de estreptavidina), y los híbridos ADN-ARN capturados se lavaron varias veces para minimizar impurezas.

4.2.2.3. Indexado

Las librerías de ADN capturadas se amplificaron mediante PCR utilizando la *Herculase II Fusion DNA Polymerase* para agregar etiquetas de indexado dual y se purificaron usando perlas de *Agencourt AMPure XP*. La calidad y cantidad de ADN se evaluó mediante el uso de *TapeStation (Agilent Technologies Inc.)*, y una tarjeta HS1000 (*Agilent Technologies Inc.*), siendo la longitud promedio de fragmento de 250-300 pb (Figura 8). Se determinó la concentración de cada librería capturada y amplificada mediante el cálculo del área bajo la curva.

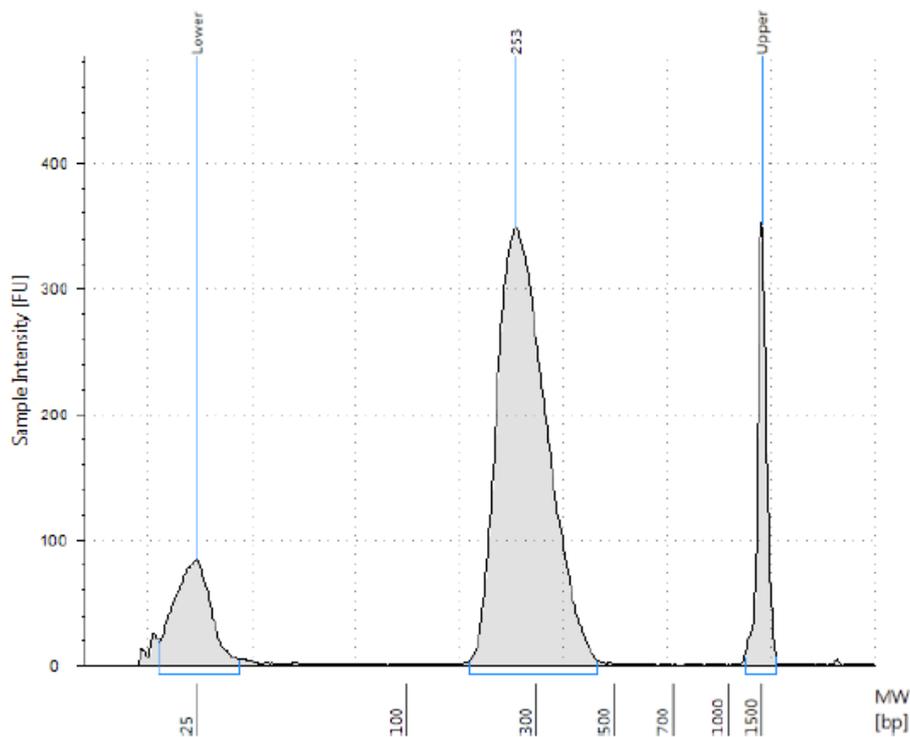


Figura 8. Análisis del ADN capturado amplificado empleando una *TapeStation 2100*.

4.2.3. Secuenciación multiplex *paired-end*

Las muestras se agruparon y se normalizaron a 4 nM para la secuenciación multiplex. Después de verificar su concentración utilizando el ensayo *Qubit dsDNA HS (ThermoFisher Scientific)*, la mezcla se diluyó a 10 pM y se desnaturalizó siguiendo el protocolo “*MiSeq System. Denature and Dilute Libraries Guide. April 2018. Protocol A.*” (Illumina Inc.). Además, los cebadores *custom* de secuenciación *SureSelect^{QXT}* se combinaron con los cebadores Illumina siguiendo las instrucciones del fabricante de *SureSelect^{QXT}*.

Se utilizó el 5% de PhiX Contro v3 como librería control. Las librerías se secuenciaron en un equipo MiSeq (Illumina Inc.), utilizando el cartucho MiSeq v3 y una longitud de lectura de 2x150 pb.

El ajuste del adaptador se realizó mediante el software MiSeq Illumina Reporter. Posteriormente, los archivos de datos se multiplexaron y se convirtieron en datos FASTQ.

4.2.4. Análisis bioinformático

El protocolo de análisis de las lecturas se compone de una fase inicial de obtención de variaciones estructurales de nucleótido único (SNP) y pequeñas inserciones/deleciones (INDELS) para cada muestra, y de una fase posterior de análisis destinada a la selección de genes diferencialmente mutados.

La fase de obtención de SNPs e INDELS comenzó con un primer filtrado de lecturas de baja calidad evaluada con FastQCv.0.11.2. Se eliminaron las lecturas de baja calidad, los duplicados de PCR y se recortaron los restos del adaptador y las bases de baja calidad utilizando cutadapt v.1.8.1. Seguidamente, se realizó un mapeo de cada lectura seleccionada contra el genoma humano de referencia (GRCh38) usando bwa mem v.07.12 (Burrows-Wheeler Aligner) y se visualizaron por IGVv2.3 (Integrative Genomics Viewer). A continuación, se realizó una estimación de variaciones estructurales (SNPs e INDELS) que cumplieren los requisitos mínimos de consistencia haplotípica y calidad definidos al inicio del procesamiento (calidad mínima de Q30 y una profundidad mínima de 10x) utilizando el software Alissa (Agilent Technologies Inc.). Este software se utiliza como fuente de anotación: 1000Genomes, 1000GenomesPhase3, CIViC, COSMIC, ClinVar, ESP6500, ExAC, HGMDProfessional, OMIM, dbNSFP y dbSNP. Como paso previo al análisis, se anotó funcionalmente cada variante. Posteriormente, las variantes se filtraron descartando las variantes con una cobertura inferior a 20x y con una frecuencia de alelo menor (MAF) superior al 1%, evitando así la inclusión de polimorfismos. La anotación de las variantes incluyó la información sobre su posición cromosómica, tipo de variante (SNP, INDEL), nucleótido de referencia, nucleótido alterado, longitud, cobertura total, cobertura de alelos menores, MAF, gen, transcripción, ubicación de la variante (exónica, intrónica, UTR 5', UTR 3'), función (cambio de pauta de lectura, sin cambio de pauta de lectura, *missense*, *nonsense*, sinónimo, *splicing*), nomenclatura variante e información de varias bases de datos y predictores biológicos.

5. RESULTADOS Y DISCUSIÓN

5.1. Diseño obtenido del panel de genes dirigido

El listado completo de 177 genes incluidos en panel de genes dirigido a NMH se adjunta en la Tabla Anexo I del ANEXO I, donde se informa también de las condiciones patológicas asociadas a cada uno de ellos extraídas de las diferentes bases de datos mencionadas en el punto 4.1.2. Además, a continuación, se presenta el listado de especificaciones concretas para determinados genes del panel que se adjuntó en su diseño, con el fin de cubrir las regiones de interés correctamente y con la profundidad adecuada. En la Tabla 3 se recogen los genes del panel que poseen pseudogenes conocidos; en la Tabla 4, los genes que presentan mutaciones intrónicas profundas; y en a Tabla 5, los genes que presentan variantes en las regiones 5' y 3'.

La importancia de la inclusión del listado de genes que presentan pseudogenes reflejado en la Tabla 3 recae en que, tal y como se ha comentado previamente, aquellos genes que presenten uno o más pseudogenes necesitan que el panel incluya un mayor número de sondas de estos. Esto es así porque muchas de las sondas mapearán en la región del/los pseudogenes, disminuyendo de este modo la cobertura final alcanzada, lo que podría dar lugar a confusión.

Las regiones intrónicas profundas de los genes que se indican en la Tabla 4 se corresponden con las regiones reguladoras de los mismos. Consideramos como intrónicas profundas aquellas mutaciones que van más allá de +/- 10 pb (ambos no inclusive, es decir, a partir de +/- 11 pb), puesto que las regiones que se encuentran a +/- 10 pb se encontrarán junto a las regiones exónicas.

Además, la Tabla 5 muestra aquellos genes de los que se ha comprobado que tienen mutaciones en las regiones 5' y 3', indicando para cada uno de ellos su identificador NM, el cromosoma donde se localizan y las regiones que ha de cubrir el panel de genes, esto último extraído de la base de datos HGMD.

Tabla 3. Listado de genes incluidos en el panel que tienen pseudogenes conocidos. Se incluye información sobre su NM, el cromosoma donde se localizan y la información extraída de la base de datos Gene del NCBI, sobre la presencia de pseudogenes y la información acerca de los mismos para cada uno. Chr: *chromosome*, del inglés, cromosoma. NM: identificador de la secuencia de referencia.

GENES DEL PANEL CON PSEUDOGENES			
Gen	RefSeq	Localización	Pseudogenes
AK2	NM_001625.3	Chr 1	Se han encontrado pseudogenes de este gen en los cromosomas 1 y 2.
ANKRD26	NM_014915.2	Chr 10	Se han encontrado pseudogenes de este gen en los cromosomas 7, 10, 13 y 16
BCOR	NM_001123383	Chr X	Se ha encontrado un pseudogen de este gen en el cromosoma Y.
BRCA1 (FANCS)	NM_007294.3	Chr 17	Se ha identificado un pseudogen relacionado que se encuentra también en el cromosoma 17.
HSPA9	NM_004134.6	Chr 5	Se ha encontrado un pseudogen de este gen en el cromosoma 2.

MCFD2	NM_001171506	Chr 2	Se ha encontrado un pseudogen de RNA procesado en el cromosoma 6p22.
MRE11A	NM_005591.3	Chr 11	Este gen tiene un pseudogen n el cromosoma 3.
NPM1	"NM_001355007, NM_002520 "	Chr 5	Se han identificado docenas de pseudogenes de este gen.
RAC2	NM_002872.4	Chr 22	Hay un pseudogen para este gen en el cromosoma 6.
RBM8A	NM_005105.4	Chr 1	Actualmente se cree que el locus RBM8A es un pseudogen.
RPL10	NM_001303624.1	Chr X	Hay múltiples pseudogenes de este gen dispersos a lo largo del genoma.
RPL11	NM_000975.4	Chr 1	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPL15	NM_001253383.2	Chr 3	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPL26	NM_000987.4	Chr 17	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPL27	NM_001349922	Chr 17	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPL35A	NM_000996.3	Chr 3	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPL5	NM_000969.4	Chr 1	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPL9	"NM_000661, NM_001024921 "	Chr 4	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS10	NM_001203245.2	Chr 6	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS14	NM_001025071.1	Chr 5	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS17	NM_001021.5 *ribosome	Chr 15	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay

			múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS19	NM_001022.3	Chr 19	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS24	NM_001142285.1	Chr 10	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS26	NM_001029.4	Chr 12	Hay multiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS27	"NM_001030, NM_001349946, NM_001349947 "	Chr1	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS29	NM_001030001.3	Chr 14	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
RPS7	NM_001011.3	Chr 2	Tal y como es típico para los genes codificantes de proteínas ribosomales, hay múltiples pseudogenes procesados de este gen dispersos a lo largo del genoma.
SBDS	NM_016038.3	Chr 7	Este gen tiene un pseudogen estrechamente vinculado que se localiza distalmente.
SRSF2	NM_003016	Chr 17	Se ha encontrado un pseudogen de este gen en el cromosoma 11.

Tabla 4. Listado de genes incluidos en el panel que poseen mutaciones intrónicas profundas que es necesario cubrir. Se indican el gen, la localización genómica hg19, variantes según la HGVS (*Human Genome Variation Society*), el identificador de la secuencia de referencia en RefSeq y el número RS. Chr: *chromosome*, del inglés, cromosoma. HGVS: *Human Genome Variation Society*. IVS: *intervening sequence*. ND: no disponible. NM: identificador de la secuencia de referencia. RS: referencia del SNP.

INTRÓNICAS PROFUNDAS DE GENES DEL PANEL				
Gen	Localización genómica hg19	HGVS	RefSeq	Número RS
ADA	ND	IVS10AS, G-A, -34	ND	ND
ADA	ND	Mutación homocigótica del intrón 11 (608958.0030) y una delección de 11 pb de las pares de bases adyacentes 31702-31712	ND	ND
ADA	ND	c.1079-15T>A	NM_000022.2	ND
ATM	Chr11:108093770	c.-174A>G	NM_000051.3	ND
ATM	Chr11:108094508	c.-31+595G>A	NM_000051.3	ND

ATM	Chr11:108098321	c.-30-1G>T	NM_000051.3	rs869312754
ATM	Chr11:108121024	c.1236-404C>T	NM_000051.3	ND
ATM	Chr11:108138753	c.2639-384A>G	NM_000051.3	ND
ATM	Chr11:108141209	c.2839-579_2839-576delAAGT	NM_000051.3	ND
ATM	Chr11:108151710	c.3403-12T>A	NM_000051.3	rs201370733
ATM	Chr11:108158168	c.3994-159A>G	NM_000051.3	rs864622543
ATM	Chr11:108179837	c.5763-1050A>G	NM_000051.3	rs774925473
ATR	ND	IVS40, C-G, +464	ND	ND
BRCA2	Chr13:32889805	c.-40+1G>A	NM_000059.3	ND
BRCA2	Chr13:32953872	c.8954-15T>G	NM_000059.3	ND
BRCA2	Chr13:32971007	c.9502-28A>G	NM_000059.3	rs397508059
BRIP1	Chr17:59858864	c.1629-498A>T	NM_032043.2	ND
DKC1	ND	IVS1, C-G, +592	ND	ND
EPCAM	ND	IVS5AS, A-G, -14	ND	ND
FANCA	Chr16:89816056	c.3239+82T>G	NM_000135.2	ND
FANCA	Chr16:89818822	c.2982-192A>G	NM_000135.2	ND
FANCA	Chr16:89831215	c.2778+83C>G	NM_000135.2	rs750997715
FANCA	Chr16:89836111	c.2504+134A>G	NM_000135.2	ND
FANCA	Chr16:89836805	c.2223-138A>G	NM_000135.2	ND
FANCA	Chr16:89849346	c.1567-20A>G	NM_000135.2	rs775154397
FANCA	Chr16:89864654	c.893+920C>A	NM_000135.2	ND
FANCC	Chr9:98011653	c.-78-2A>G	NM_000136.2	rs587779898
FANCD2	Chr3:10083186	c.696-121C>G	NM_033084.3	ND
FANCI	Chr15:89825208	c.1583+142C>T	NM_001113378.1	ND
PALB2	Chr16:23649285	c.109-12T>A	NM_024675.3	rs774949203
TAZ	ND	IVS3DS, G-A, +110	ND	ND
TRNT1	ND	IVS5AS, T-C, -26	ND	ND

Tabla 5. Listado de genes incluidos en el panel que poseen mutaciones en las regiones reguladoras. Se incluye el nombre del gen, su NM, su localización cromosómica y las regiones a cubrir en cada uno de ellos. Chr: *chromosome*, del inglés, cromosoma. NM: identificador de la secuencia de referencia. UTR: *untranslated region*.

GENES CON MUTACIONES PATOLÓGICAS EN REGIONES 5' Y 3'			
Gen	RefSeq	Localización	Regiones a cubrir
ALAS2	NM_000032.4	Chr X	Región codificante + 5'UTR + 3'UTR
ANKRD26	NM_014915.2	Chr 10	Región codificante + 5'UTR + 3'UTR
ATM	NM_000051.3	Chr 11	Región codificante + 5'UTR + 3'UTR
BRCA1 (FANCS)	NM_007294.3	Chr 17	Región codificante + 5'UTR + 3'UTR
DKC1	NM_001363.4	Chr X	Región codificante + 5'UTR + 3'UTR
FANCD1/BRCA2	NM_000059.3	Chr 13	Región codificante + 5'UTR + 3'UTR
FANCD2	NM_033084.4	Chr 3	Región codificante + 5'UTR + 3'UTR
FANCN/PALB2	NM_024675.3	Chr 16	Región codificante + 5'UTR + 3'UTR
FANCO/RAD51C	NM_058216.2	Chr 17	Región codificante + 5'UTR + 3'UTR
FANCO/ERCC4	NM_005236.2	Chr 16	Región codificante + 5'UTR + 3'UTR
GATA2	NM_032638.4	Chr 3	Región codificante + 5'UTR + 3'UTR
KIF23	NM_138555.3	Chr 15	Región codificante + 5'UTR + 3'UTR
LAMTOR2	NM_014017.3	Chr 1	Región codificante + 5'UTR + 3'UTR

MAD2L2	NM_001127325.1	Chr 1	Región codificante + 5'UTR + 3'UTR
MLH1	NM_000249.3	Chr 3	Región codificante + 5'UTR + 3'UTR
MSH2	NM_000251.2	Chr 2	Región codificante + 5'UTR + 3'UTR
RBM8A	NM_005105.4	Chr 1	Región codificante + 5'UTR + 3'UTR
RUNX1	NM_001754.4	Chr 21	Región codificante + 5'UTR + 3'UTR
SRP72	NM_006947.3	Chr 4	Región codificante + 5'UTR + 3'UTR
TERC	NR_001566.1	Chr 3	Región codificante + 5'UTR + 3'UTR
TERT	NM_198253.2	Chr 5	Región codificante + 5'UTR + 3'UTR
TP53	NM_000546.5	Chr 17	Región codificante + 5'UTR + 3'UTR

Por último, también cabe destacar cuáles fueron las especificaciones finales del panel, que presentó 400.457 sondas diferentes con un tamaño de 759.16 kb (hg19) correspondiendo a un diseño de tipo Tier2 según el programa *SureDesign Custom Design Tool* (Agilent Technologies Inc.).

5.2. Resultados de la validación del panel de genes dirigido

5.2.1. Cobertura

En los ensayos de NGS, es necesario establecer una profundidad de cobertura mínima para realizar el *variant calling* (asignaciones de variantes) de manera fiable. En nuestro caso, como ya hemos dicho, utilizamos muestras de líneas celulares previamente caracterizadas, pudiendo de este modo definir la profundidad de cobertura alcanzada, que se muestra en la Tabla 6.

Tabla 6. Valores de cobertura obtenidos para cada uno de los ensayos llevados a cabo para cada línea celular, en base 1.

Muestra	Total de lecturas	1X	10X	20X	50X	100X
D1M1	79.207.867	0,9991	0,9741	0,9091	0,6375	0,2453
D1M2	82.635.060	0,9995	0,9841	0,9246	0,6642	0,2690
D1M3	85.878.586	0,9997	0,9780	0,9147	0,6599	0,2944
K12M1	88.106.112	0,9993	0,9832	0,9353	0,6879	0,3147
K12M2	94.416.195	0,9993	0,9909	0,9526	0,7284	0,3556
K12M3	94.070.199	0,9996	0,9853	0,9422	0,7164	0,3565

La profundidad de lectura mínima fue del 20X, que según las guías clínicas de buenas prácticas de NGS es la adecuada para considerar una variante *missense* como real. En la Tabla 6, los valores de cobertura obtenidos no fueron todo lo buenos que cabría esperar. El resultado ideal hubiera sido obtener valores de cobertura del 95% (0,95) tal y como se expresa en la Tabla 6 o superiores. En nuestro caso, dichos valores sí que se obtuvieron a cobertura tanto de 1X como de 10X, pero en el caso de la cobertura a 20X, esto dejó de cumplirse para todos los casos, llegando al umbral establecido únicamente la muestra 2 de la línea celular K12 (K12M2).

A pesar de ello, los valores de cobertura obtenidos a 20X fueron mayores del 90% (0,90) para todos los casos, por lo que a pesar de no alcanzar el umbral establecido del 95%, tampoco consideramos que el ensayo no fuese válido, aunque se podría mejorar, ya que se necesitaría un mayor porcentaje de lecturas para alcanzar los valores de confianza ideales.

Nuestro panel NGS proporcionó una profundidad de secuenciación media de 80X por muestra. El 99% de las regiones objetivo se cubrieron con una cobertura mínima de 70X en cada gen.

En el ensayo realizado se cargaron un total de 24 librerías en un cartucho MiSeq V3, dado el tamaño del panel, realizando una carga de menos librerías se puede llegar a conseguir un grado de cobertura mejor.

5.2.2. Precisión

En cuanto a la precisión de los ensayos llevados a cabo, nos vamos a centrar en la comparativa entre los resultados obtenidos para cada línea celular en los distintos *runs* realizados.

5.2.2.1. Línea celular Coriell D1 (NA12144)

Tras la realización de los ensayos, se obtuvieron los siguientes resultados para la línea celular D1. En la Tabla 7 se presenta el resumen inicial del estudio realizado.

Tabla 7. Resumen de la información obtenida tras el ensayo realizado con la línea celular Coriell D1, en que se indica para cada uno de los ensayos (M1, M2 y M3), el número de posiciones detectadas, los falsos positivos, el porcentaje de error y el porcentaje de verdaderos positivos, además de la media para todos los ensayos de los últimos dos.

	M1	M2	M3	MEDIA
Posiciones detectadas	509	519	503	-
Falsos positivos	35	45	29	-
% Error	6,88	8,67	5,77	7,10
% Verdaderos positivos	93,12	91,33	94,23	92,90

El número de posiciones detectadas que se ve reflejado en la tabla está relacionado con el sumatorio de SNPs, inserciones y deleciones que fueron detectados en cada uno de los ensayos. Por otro lado, el número de falsos positivos se obtuvo comprobando cuántas de las posiciones detectadas en el ensayo, habían sido detectadas también en el resto de los ensayos. De este modo, se han considerado como verdaderas positivas únicamente aquellas posiciones detectadas en los tres ensayos llevados a cabo con la misma línea celular.

Por tanto, tal y como se deduce de los datos reflejados en la Tabla 7, se concluye que el número de posiciones verdaderas detectadas, es decir, polimorfismos detectados en las tres repeticiones, es de 474.

Además, se calculó también el porcentaje de error, obteniendo a su vez el porcentaje de verdaderos positivos. Tal y como presenta la Tabla 7, el porcentaje de error fue del 7,10 %, relativamente bajo, pero se trata de una cifra que todavía se podría mejorar. Esto deja un porcentaje de verdaderos positivos del 92,90 % para la línea celular Coriell D1.

Por otro lado, también se comprobó para cada una de las 474 posiciones verdaderas, que tipo de variante había sido detectada, resultado de lo cual se muestra a continuación en la Tabla 8.

Asimismo, se analizó el tipo de variante detectada en las 474 posiciones verdaderas, y tal y como se puede ver en la Tabla 8, la variante más predominante detectada se trató del SNP, abarcando un 89,24 %, seguido de las deleciones, con un 6,75 % del total y, por último, las inserciones, encontrándose únicamente en un 4,01 % del total.

Tabla 8. Tipo de variantes detectadas en la línea celular Coriell D1 en absoluto y porcentaje de cada tipo sobre el total.

	Posiciones detectadas	% Sobre el total
SNP	423	89,24
Inserciones	19	4,01
Deleciones	32	6,75

Además, también se calculó la tasa de error para cada uno de los ensayos llevados a cabo con la misma línea celular y para cada tipo de variante por separado, teniendo en cuenta para el total de posiciones detectadas en cada uno de los ensayos, cuántas de ellas eran falsos negativos de cada tipo de variante. Los resultados se muestran en la Tabla 9. Y la tasa de error es mucho mayor en inserciones y deleciones que en SNPs, dato que concuerda con lo esperado, pues es la tasa de error inherente a la técnica es mayor en INDELS que en SNPs.

Tabla 9. Tasa de error por tipo de variante para cada uno de los 3 ensayos realizados con la línea celular Coriell D1.

	TASA DE ERROR (%)		
	SNP	Inserciones	Deleciones
M1	5,16	20,83	17,95
M2	5,99	14,29	31,91
M3	3,20	20,83	21,43
PROMEDIO	4,78	18,65	23,76

El promedio de la tasa de error obtenida para la línea celular Coriell D1 es del 4,78 % para SNPs, y del 21,21 % para INDELS (INDELS = inserciones y deleciones). Se ha simplificado la tasa de error de inserciones y deleciones agrupando ambos tipos de variantes como un único tipo, los INDELS, obteniéndose su tasa de error con el promedio de las tasas de error de inserciones y deleciones.

5.2.2.2. Línea Celular Coriell K12 (NA12878)

En el caso de la línea celular Coriell K12 se siguió el mismo procedimiento de análisis de datos, obteniéndose los resultados del procedimiento que se resumen y analizan a continuación en la Tabla 10.

Se ha reflejado el número total de posiciones detectadas en cada ensayo englobando todos los tipos de variantes analizadas: SNPs, inserciones y deleciones. Y del mismo modo, se ha realizado el cálculo de los falsos positivos teniendo en cuenta cuántas de las posiciones detectadas lo habían sido en los tres ensayos. Esas posiciones constituirían los verdaderos positivos, que han resultado en 497 posiciones verdaderas. Y los falsos negativos se han utilizado para calcular el porcentaje de error para cada una de las tres repeticiones llevadas a cabo con la línea celular Coriell K12.

Con estos datos, se calculó la media del porcentaje de error de las tres repeticiones, siendo del 6,90%, un valor menor que en el caso de la línea celular Coriell D1. Y, por el contrario, el valor de verdaderos positivos obtenido fue mayor, siendo del 93,10 %.

Tabla 10. Resumen de la información obtenida tras el ensayo realizado con la línea celular Coriell K12, en que se indica para cada uno de los ensayos (M1, M2 y M3), el número de posiciones detectadas, los falsos positivos, el porcentaje de error y el porcentaje de verdaderos positivos, además de la media para todos los ensayos de los últimos dos.

	M1	M2	M3	MEDIA
Posiciones detectadas	522	538	542	-
Falsos positivos	25	41	45	-
% Error	4,79	7,62	8,30	6,90
% Verdaderos positivos	95,21	92,38	91,70	93,10

Por tanto, se deduce que los ensayos realizados con la línea celular Coriell K12 han tenido un menor porcentaje de error, pero a pesar de ello, la diferencia con la Coriell D1 es muy pequeña.

Por otro lado, al igual que se hizo con la línea celular Coriell D1, se comprobó también para cada una de las 497 posiciones verdaderas detectadas, qué tipo de polimorfismo había sido detectado, resultado de lo cual se muestra a continuación en la Tabla 11, donde la variante más predominante detectada se trató del SNP, contando con el 89,24 % de las posiciones totales detectadas. A continuación, le siguieron las deleciones, con un 8,05 % del total y, por último, las inserciones, que representaron el 4,83 % del total.

Tabla 11. Tipo de polimorfismo de las posiciones detectadas en la línea celular Coriell K12 en absoluto y porcentaje de cada tipo sobre el total.

	Posiciones detectadas	% Sobre el total
SNP	433	87,12
Inserciones	24	4,83
Deleciones	40	8,05

Además, calculamos también la tasa de error para cada uno de los ensayos con la misma línea celular y para cada tipo de variante por separado del mismo modo que en el caso de la línea celular Coriell D1, obteniendo los resultados reflejados en la Tabla 12.

En este caso, podemos comprobar que del mismo modo que en el caso anterior, la tasa de error obtenida para INDELS es mucho mayor que para SNPs, lo que sigue concordando con lo esperado, siendo un hecho inherente a la técnica.

Tabla 12. Tasa de error por tipo de variante para cada uno de los 3 ensayos realizados con la línea celular Coriell K12.

	TASA DE ERROR %		
	SNP	Inserciones	Deleciones
M1	2,91	7,69	20,00
M2	5,25	8,33	26,32
M3	5,68	21,43	25,00
PROMEDIO	4,61	12,48	23,77

El promedio de la tasa de error obtenida para la línea celular Coriell K12 es del 5,18 % para SNPs y del 18,13 % para INDELS. (INDELS = inserciones y deleciones). Al igual que se ha hecho anteriormente, se ha simplificado la tasa de error de inserciones y deleciones agrupando ambos tipos de variantes como un único tipo, los INDELS, obteniéndose su tasa de error con el promedio de las tasas de error de inserciones y deleciones.

5.2.2.3. *Comparación de las tasas de error obtenidas para cada línea celular*

Una vez obtenidos los datos para ambas líneas celulares, lo que se hizo fue hacer un promedio de los mismos con el objetivo de establecer la tasa de error media del panel de genes, tratando de minimizar el margen de error. Se obtuvieron los datos diferenciando la tasa de error media para SNPs y la tasa de error media para INDELS.

El resultado del promedio de la tasa de error para las dos líneas celulares empleadas en el estudio (Coriell D1 y Coriell K12) fue del 4,98 % en el caso de SNPs y del 19,67 % para los INDELS. Tal y como se ha comentado anteriormente, la tasa de error media es mucho menor en el caso de los SNPs que en el caso de INDELS (inserciones y deleciones).

5.2.2.4. *Genes más polimórficos*

Por otro lado, se quiso también ver qué genes de los incluidos en el panel presentaban un mayor número de variantes en casa caso (Tabla 13), donde, además, se puede comprobar cuáles fueron los más polimórficos en cada una de las líneas celulares.

Se encuentran coloreados aquellos que presentaron un mayor número de variantes tanto para la línea celular Coriell D1 y la Coriell K12, teniendo en cuenta solo los verdaderos positivos. Los genes que coincidieron en ambos casos fueron los que consideramos como genes más polimórficos que incluía el panel. Se consideraron como genes más polimórficos aquellos que poseían 10 o más variantes por gen, con tal de establecer un criterio unificador.

Tabla 13. Genes más polimórficos para las líneas celulares Coriell D1 y K12. Se destacan aquellos genes que coinciden en ambas líneas celulares.

Genes más polimórficos en la línea celular Coriell D1	Genes más polimórficos en la línea celular Coriell K12
-	<i>ANKRD26</i>
<i>ATG2B</i>	<i>ATG2B</i>
-	<i>BRCA1</i>
<i>CUX1</i>	<i>CUX1</i>
<i>ERCC4</i>	<i>ERCC4</i>
<i>FANCA</i>	<i>FANCA</i>
<i>FANCD2</i>	<i>FANCD2</i>
-	<i>FANCI</i>
<i>FLT3</i>	<i>FLT3</i>
<i>NF1</i>	-
-	<i>RAC2</i>
<i>RTEL1-TNFRSF6B</i>	<i>RTEL1-TNFRSF6B</i>
-	<i>TRNT1</i>
<i>WRN</i>	<i>WRN</i>

Tal y como se puede ver, no todos estos genes coinciden en los resultados obtenidos para cada línea celular, por lo que se consideran como genes más polimórficos aquellos que tienen 10 o más variantes por gen en ambas líneas celulares, que son los 8 genes que se encuentran sombreados en la tabla: *ATG2B*, *CUX1*, *ERCC4*, *FANCA*, *FANCD2*, *FLT3*, *RTEL1-TNFRSF6B* Y *WRN*.

5.2.2.5. Genes polimórficos relacionados con Neoplasias Mieloides Hereditarias, anemia de Fanconi y enfermedad de Blackfan-Diamond

Otro de los enfoques que se ha decidido llevar a cabo en el análisis de los resultados ha constado de la extracción de información sobre qué tipo de enfermedades estaban relacionadas con aquellos genes que presentan polimorfismos, centrándose no solo en las NMH, sino también en la anemia de Fanconi y la enfermedad de Blackfan-Diamond, dos trastornos muy relacionados con las NMH y que, por tanto, se suelen englobar indirectamente también en este tipo de estudios. Las siguientes tablas se desglosan por enfermedades, y dentro de cada enfermedad, por línea celular, los genes relacionados con las mismas (Godley & Shimamura, 2017) que han presentado variantes.

5.2.2.5.1. Neoplasias Mieloides Hereditarias

Tal y como muestra la Tabla 14, a continuación, se pueden ver aquellos genes y el número de variantes encontradas en cada uno de ellos para cada línea celular con que se ha llevado a cabo el ensayo, la línea celular Coriell D1 y la Coriell K12.

Tabla 14. Genes polimórficos asociados a las NMH y número de variantes detectadas en cada uno para las líneas celulares Coriell D1 y Coriell K12.

CORIELL D1		CORIELL K12	
Gen	Variantes por gen	Gen	Variantes/gen
<i>ANKRD26</i>	4	<i>ANKRD26</i>	11
<i>ATG2B</i>	12	<i>ATG2B</i>	11
<i>CEBPA</i>	1	<i>CEBPA</i>	1
<i>DDX41</i>	4	<i>DDX41</i>	5
<i>ETV6</i>	1	<i>ETV6</i>	1
<i>GATA2</i>	3	<i>GATA2</i>	3
<i>GSKIP</i>	1	<i>GSKIP</i>	1
<i>PAX5</i>	2	<i>PAX5</i>	1
<i>RUNX1</i>	5	<i>RUNX1</i>	4
<i>SH2B3</i>	1	<i>SH2B3</i>	2
<i>SRP72</i>	3	<i>SRP72</i>	2
<i>TERT</i>	3	<i>TERT</i>	1

Como se puede ver, los genes más polimórficos detectados en cada línea celular resultaron ser los mismos para el caso de las NMH, difiriendo únicamente en el número de variantes que presentaba cada uno de ellos. La mayor diferencia se encontró en el gen *ANKRD26*, que en la línea celular Coriell D1 presentó 4 variantes por gen, mientras que en la Coriell K12 se han encontrado 11 variantes. El resto de los genes presenta un número de variantes más similar

5.2.2.5.2. Anemia de Fanconi

La anemia de Fanconi se trata de un trastorno hereditario relacionado con la reparación del ADN, que se caracteriza por malformaciones congénitas diversas, poseer predisposición a desarrollar tumores e insuficiencia de la médula ósea. (ORPHANET, S.F.)

En la Tabla 15, se puede ver, del mismo modo que en el caso anterior, qué genes relacionados con la enfermedad han presentado variantes y qué número de ellas presentaba cada uno para cada línea celular.

Tabla 15. Genes polimórficos asociados a la anemia de Fanconi y número de variantes detectadas en cada uno para las líneas celulares Coriell D1 y Coriell K12.

CORIELL D1		CORIELL K12	
Gen	Variantes por gen	Gen	Variantes por gen
<i>BRCA1</i>	1	<i>BRCA1</i>	13
<i>BRCA2</i>	6	<i>BRCA2</i>	6
<i>BRIP1</i>	3	<i>BRIP1</i>	4
<i>ERCC4</i>	19	<i>ERCC4</i>	19
<i>FANCA</i>	19	<i>FANCA</i>	17
<i>FANCC</i>	2	-	-
<i>FANCD2</i>	23	<i>FANCD2</i>	22
<i>FANCE</i>	2	<i>FANCE</i>	2
<i>FANCG</i>	2	<i>FANCG</i>	2
<i>FANCI</i>	1	<i>FANCI</i>	10
<i>FANCL</i>	3	<i>FANCL</i>	2
<i>FANCM</i>	1	<i>FANCM</i>	1
<i>MAD2L2</i>	5	<i>MAD2L2</i>	7
-	-	<i>PALB2</i>	2
<i>RAD51</i>	1	<i>RAD51</i>	1
<i>RAD51C</i>	1	<i>RAD51C</i>	2
<i>SLX4</i>	1	<i>SLX4</i>	2
<i>UBE2T</i>	1	-	-

En este caso, se puede deducir que no todos los genes coinciden en ambas líneas celulares, difiriendo en los genes *FANCC* y *UBE2T*, que presentaron 2 y 1 variante respectivamente en la línea celular Coriell D1, y *PALB2*, que presentó 2 variantes en la línea celular Coriell K12. En cuanto a número de variantes, estos genes no fueron los más relevantes, destacando los genes *ERCC4*, *FANCA* y *FANCD2* como aquellos con un mayor número de variantes en ambas líneas celulares.

También cabe mencionar los genes *BRCA1* y *FANCI*, pues presentaron una diferencia significativa en cuando al número de variantes localizadas en cada uno para cada línea celular, siendo en el caso de *BRCA1* de 1 variante en la línea celular Coriell D1 frente a 13 en la línea celular Coriell K12. Y lo mismo sucedió con *FANCI*, con 1 variante detectada en la línea celular Coriell D1 frente a las 10 variantes detectadas en la K12.

5.2.2.5.3. Enfermedad de Blackfan-Diamond

La enfermedad de Blackfan-Diamond se trata de una anemia congénita (ORPHANET, S.F.) poco frecuente donde la médula ósea no produce suficientes glóbulos rojos, y se asocia con un aumento del riesgo de leucemia y sarcoma, especialmente osteosarcoma. (NATIONAL CANCER INSTITUTE - DICCIONARIO DE CÁNCER, S.F.)

Los genes relacionados con la enfermedad que han presentado variantes se muestran a continuación en la Tabla 16.

Tabla 16. Genes polimórficos asociados a la anemia de Fanconi y número de variantes detectadas en cada uno para las líneas celulares Coriell D1 y Coriell K12.

CORIELL D1		CORIELL K12	
Gen	Variantes por gen	Gen	Variantes por gen
<i>RPL11</i>	2	<i>RPL11</i>	1
-	-	<i>RPL27</i>	1
<i>RPL15</i>	1	-	-
<i>RPL27</i>	1	-	-
-	-	<i>RPL9</i>	1
-	-	<i>RPS17</i>	3
<i>RPS19</i>	1	<i>RPS19</i>	1
<i>RPS24</i>	1	<i>RPS24</i>	2
<i>RPS26</i>	1	<i>RPS26</i>	1
-	-	<i>RPS27</i>	1
<i>RPS7</i>	1	<i>RPS7</i>	1
<i>SBDS</i>	1	-	-

En este caso, se vio un menor número de genes que presentaban variantes coincidentes para ambas líneas celulares, pero cabe destacar también que la mayoría de los genes para ambos casos fueron de las familias *RPL* y *RPS*. El número de variantes detectadas en este caso por cada gen fue menor que en los casos anteriores, presentando todos entre 1 y 3 variantes como máximo, por lo que no se destacó la detección de ningún gen especialmente polimórfico.

5.2.3. Zonas de mejora

Tras el análisis de calidad y validación del panel, se obtuvo un listado de posiciones que no se habían logrado cubrir en su totalidad. Además, a partir del listado de zonas de mala calidad obtenido en cada uno de los experimentos, se obtuvieron conclusiones sobre qué zonas estaban mal cubiertas y se deberían mejorar en la próxima versión del panel de genes dirigido.

Para ello, se decidió establecer un criterio de selección, tratando de minimizar los artefactos, marcando como posiciones de mejora aquellas que se encontraban mal cubiertas en al menos tres de las muestras (es decir, la mitad de las muestras), con un porcentaje sin cubrir mayor del 50 %. Así, se establecieron diversas posiciones cromosómicas como zonas de mejora que se decidió agrupar en función de los genes a que pertenecían, tal y como se muestra en la Tabla Anexo II del ANEXO II, con el objetivo de tener en cuenta para el diseño de la próxima versión del panel que dichos genes precisarán de un mayor volumen de sondas para tratar de alcanzar la cobertura deseada.

6. CONCLUSIONES

- La presencia de variantes de línea germinal en genes que causan desórdenes relacionados con SMD y LMA es mayor de la esperada y puede ayudar en el diagnóstico y futuro tratamiento de dichas enfermedades.
- La detección de mutaciones patogénicas en línea germinal, además del estudio de la historia clínica del paciente y de sus familiares, puede ser de gran utilidad clínica para el diagnóstico de las Neoplasias Mieloides.
- La puesta a punto de los paneles de genes por NGS requiere de un laborioso trabajo tanto de diseño como de posterior validación de la técnica de diagnóstico.
- Es imprescindible seguir las guías ACMG para poder realizar una profunda validación analítica de paneles de genes diseñados de forma *custom* que asegure la elevada calidad de los resultados de secuenciación que se implementan en el ámbito clínico.
- Las líneas celulares Coriell D1 y Coriell K12 son adecuadas para realizar un análisis del panel de genes diseñado, ya que se obtienen valores adecuados y similares de sensibilidad y especificidad.
- Se ha comprobado que la tasa de error inherente a la técnica es mayor en INDELS que en SNPs, con lo que el panel tiene mayor sensibilidad para la detección de SNPs.
- Se detectan ciertos genes con altas tasas de polimorfismo para las patologías estudiadas.
- Se han detectado regiones mal cubiertas que se deben tener en cuenta para su mejora en la próxima versión del panel.
- La técnica presenta limitaciones que habrá que tener en consideración para su utilización como herramienta de diagnóstico y que se deben tratar de mejorar en futuras versiones del panel.

7. BIBLIOGRAFÍA

- AGILENT TECHNOLOGIES. (2015). *Automated Target Enrichment for Illumina Multiplexed Sequencing Automated using Agilent NGS Bravo*. Visto el 25 de Junio de 2019. <https://www.agilent.com/cs/library/usermanuals/public/G9681-90020.pdf>
- AKPAN, I. J., OSMAN, A. E. G., DRAZER, M. W., & GODLEY, L. A. (2018). Hereditary Myelodysplastic Syndrome and Acute Myeloid Leukemia: Diagnosis, Questions, and Controversies. *Current Hematologic Malignancy Reports*, 13(6), 426–434. <https://doi.org/10.1007/s11899-018-0473-7>
- AMERICAN COLLEGE OF MEDICAL GENETICS AND GENOMICS (ACMG). (S.F.). Visto 20 de Junio, 2019, de <https://www.acmg.net/>
- ANTONY-DEBRÉ, I., DUPLOYEZ, N., BUCCI, M., GEFFROY, S., MICOL, J.-B., RENNEVILLE, A., ... PREUDHOMME, C. (2016). Somatic mutations associated with leukemic progression of familial platelet disorder with predisposition to acute myeloid leukemia. *Leukemia*, 30(4), 999–1002. <https://doi.org/10.1038/leu.2015.236>
- BABUSHOK, D. V., BESSLER, M., & OLSON, T. S. (2016). Genetic predisposition to myelodysplastic syndrome and acute myeloid leukemia in children and young adults. *Leukemia and Lymphoma*, 57(3), 520–536. <https://doi.org/10.3109/10428194.2015.1115041>
- BANNON, S. A., & DINARDO, C. D. (2016). Hereditary predispositions to myelodysplastic syndrome. *International Journal of Molecular Sciences*, 17(6). <https://doi.org/10.3390/ijms17060838>
- CHURPEK, J. E., & GODLEY, L. A. (2016). How I diagnose and manage individuals at risk for inherited myeloid malignancies. *Blood*, 128(14), 1800–1813. <https://doi.org/10.1182/blood-2016-05-670240>
- DESAI, A. V., PERPICH, M., & GODLEY, L. A. (2017). Clinical Assessment and Diagnosis of Germline Predisposition to Hematopoietic Malignancies: The University of Chicago Experience. *Frontiers in Pediatrics*, 5(December), 1–10. <https://doi.org/10.3389/fped.2017.00252>
- DINARDO, C. D., ROUTBORT, M. J., BANNON, S. A., BENTON, C. B., TAKAHASHI, K., KORNBLAU, S. M., ... PATEL, K. P. (2018). Improving the detection of patients with inherited predispositions to hematologic malignancies using next-generation sequencing-based leukemia prognostication panels. *Cancer*, 124(13), 2704–2713. <https://doi.org/10.1002/cncr.31331>
- DRAZER, M. W., KADRI, S., SUKHANOVA, M., PATIL, S. A., WEST, A. H., FEURSTEIN, S., ... GODLEY, L. A. (2018). Prognostic tumor sequencing panels frequently identify germ line variants associated with hereditary hematopoietic malignancies. *Blood Advances*, 2(2), 146–150. <https://doi.org/10.1182/bloodadvances.2017013037>
- EUROGENTEST. (S.F.). Visto el 22 de Junio, 2019, de <http://www.eurogentest.org/index.php?id=160>
- GARCÍA, W. R., SÁEZ, S. S., CARLOS, J., & FAGUNDO, J. (2018). *Síndromes de hemopatías mieloides malignas hereditarias*. *Inherited myeloid malignancy syndromes*. 34(1), 116–124.

- GODLEY, L. A. (2014). Inherited Predisposition to Acute Myeloid Leukemia. *Seminars in Hematology*, 51(4), 306–321. <https://doi.org/10.1053/j.seminhematol.2014.08.001>
- GODLEY, L. A., & SHIMAMURA, A. (2017). Genetic predisposition to hematologic malignancies: Management and surveillance. *Blood*, 130(4), 424–432. <https://doi.org/10.1182/blood-2017-02-735290>
- HAHN, C. N., CHONG, C. E., CARMICHAEL, C. L., WILKINS, E. J., BRAUTIGAN, P. J., LI, X. C., ... SCOTT, H. S. (2011). Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nature Genetics*, 43(10), 1012–1019. <https://doi.org/10.1038/ng.913>
- HAYASHI, Y., HARADA, Y., HUANG, G., & HARADA, H. (2017). Myeloid neoplasms with germ line RUNX1 mutation. *International Journal of Hematology*, 106(2), 183–188. <https://doi.org/10.1007/s12185-017-2258-5>
- HSU, A. P., JOHNSON, K. D., SANALKUMAR, R., SANCHEZ, L. A., HICKSTEIN, D. D., CUELLAR-RODRIGUEZ, J., ... HOLLAND, S. M. (2012). GATA2 haploinsufficiency is caused by mutations in a conserved intronic element leading to MonoMAC syndrome. *Blood*, 121(19), 3830–3838. <https://doi.org/10.1182/blood-2012-08-452763>.A.P.H.
- JALAGADUGULA, G., MAO, G., KAUR, G., GOLDFINGER, L. E., DHANASEKARAN, D. N., & RAO, A. K. (2010). Regulation of platelet myosin light chain (MYL9) by RUNX1: Implications for thrombocytopenia and platelet dysfunction in RUNX1 haploinsufficiency. *Blood*, 116(26), 6037–6045. <https://doi.org/10.1182/blood-2010-06-289850>
- KIRÁLY, A. P., KÁLLAY, K., GÁNGÓ, A., KELLNER, Á., EGYED, M., SZÓKE, A., ... BÖDÖR, C. (2018). Familial Acute Myeloid Leukemia and Myelodysplasia in Hungary. *Pathology and Oncology Research*, 24(1), 83–88. <https://doi.org/10.1007/s12253-017-0216-4>
- KIRWAN, M., WALNE, A. J., PLAGNOL, V., VELANGI, M., HO, A., HOSSAIN, U., ... DOKAL, I. (2012). Exome sequencing identifies autosomal-dominant SRP72 mutations associated with familial aplasia and myelodysplasia. *American Journal of Human Genetics*, 90(5), 888–892. <https://doi.org/10.1016/j.ajhg.2012.03.020>
- LEWINSOHN, M., BROWN, A. L., WEINEL, L. M., PHUNG, C., RAFIDI, G., LEE, M. K., ... SCOTT, H. S. (2016). Novel germ line DDX41 mutations define families with a lower age of MDS/AML onset and lymphoid malignancies. *Blood*, 127(8), 1017–1023. <https://doi.org/10.1182/blood-2015-10-676098>
- LIEW, E., & OWEN, C. (2011). Familial myelodysplastic syndromes: A review of the literature. *Haematologica*, 96(10), 1536–1542. <https://doi.org/10.3324/haematol.2011.043422>
- LORDIER, L., BLUTEAU, D., JALIL, A., LEGRAND, C., PAN, J., RAMEAU, P., ... CHANG, Y. (2012). RUNX1-induced silencing of non-muscle myosin heavy chain IIB contributes to megakaryocyte polyploidization. *Nature Communications*, 3. <https://doi.org/10.1038/ncomms1704>
- NATIONAL CANCER INSTITUTE - Diccionario de cáncer - Definición de anemia de Diamond-Blackfan. (S.F.). Visto 19 de Junio, 2019, from <https://www.cancer.gov/espanol/publicaciones/diccionario/def/anemia-de-diamond-blackfan>

- NERLOV, C. (2007). The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends in Cell Biology*, 17(7), 318–324. <https://doi.org/10.1016/j.tcb.2007.07.004>
- NISHIMOTO, N., IMAI, Y., UEDA, K., NAKAGAWA, M., SHINOHARA, A., ICHIKAWA, M., ... KUROKAWA, M. (2010). T cell acute lymphoblastic leukemia arising from familial platelet disorder. *International Journal of Hematology*, 92(1), 194–197. <https://doi.org/10.1007/s12185-010-0612-y>
- NOETZLI, L., LO, R. W., LEE-SHERICK, A. B., CALLAGHAN, M., NORIS, P., SAVOIA, A., ... DI PAOLA, J. (2015). Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. *Nature Genetics*, 47(5), 535–538. <https://doi.org/10.1038/ng.3253>
- OK, C. Y., LEVENTAKI, V., WANG, S. A., DINARDO, C., MEDEIROS, L. J., & KONOPLEV, S. (2016). Detection of an abnormal myeloid clone by flow cytometry in familial platelet disorder with propensity to myeloid malignancy. *American Journal of Clinical Pathology*, 145(2), 271–276. <https://doi.org/10.1093/AJCP/AQV080>
- ORPHANET: Anemia de Fanconi. (S.F.). Visto el 22 de Junio, 2019, de https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Lng=ES&Expert=84
- ORPHANET: Enfermedad de Blackfan Diamond. (S.F.). Visto el 22 de Junio, 2019, de https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Lng=ES&Expert=124
- PASQUET, M., BELLANNÉ-CHANTELLOT, C., TAVITIAN, S., PRADE, N., BEAUPAIN, B., LAROCHELLE, O., ... DELABESSE, E. (2013). High frequency of GATA2 mutations in patients with mild chronic neutropenia evolving to MonoMac syndrome, myelodysplasia, and acute myeloid leukemia. *Blood*, 121(5), 822–829. <https://doi.org/10.1182/blood-2012-08-447367>
- POLPRASERT, C., SCHULZE, I., SEKERES, M. A., MAKISHIMA, H., PRZYCHODZEN, B., HOSONO, N., ... MACIEJEWSKI, J. P. (2015). Inherited and Somatic Defects in DDX41 in Myeloid Neoplasms. *Cancer Cell*, 27(5), 658–670. <https://doi.org/10.1016/j.ccell.2015.03.017>
- PORTER, C. C. (2016). Germ line mutations associated with Leukemias. *Hematology*, 2016(1), 302–308. <https://doi.org/10.1182/asheducation-2016.1.302>
- QUBIT 4 FLUOROMETER - ES. (S.F.). Visto el 15 de Junio, 2019, de <https://www.thermofisher.com/es/es/home/industrial/spectroscopy-elemental-isotope-analysis/molecular-spectroscopy/fluorometers/qubit/qubit-fluorometer.html>
- SAKURAI, M., KASAHARA, H., YOSHIDA, K., YOSHIMI, A., KUNIMOTO, H., WATANABE, N., ... NAKAJIMA, H. (2016). Genetic basis of myeloid transformation in familial platelet disorder/acute myeloid leukemia patients with haploinsufficient RUNX1 allele. *Blood Cancer Journal*, 6(2), e392–e392. <https://doi.org/10.1038/bcj.2015.81>
- SCHNITTGER, S., DICKER, F., KERN, W., WENDLAND, N., SUNDERMANN, J., ALPERMANN, T., ... TORSTEN HAFERLACH. (2011). RUNX1 mutations are frequent in de novoAMLwith noncomplex karyotype and confer an unfavorable prognosis. *Blood Journal*, 117(8), 2348–2357. <https://doi.org/10.1182/blood-2009-11-255976>
- SPINNER, M. A., SANCHEZ, L. A., HSU, A. P., SHAW, P. A., ZERBE, C. S., CALVO, K. R., ARTHUR, D. C., GU, W., GOULD, C. M., BREWER, C. C., COWEN, E. W., FREEMAN, A. F., OLIVIER, K. N.,

- UZEL, G., ZELAZNY, A. M., DAUB, J. R., SPALDING, C. D., CLAYPOOL, R. J., GIRI, N. K., ALTER, B. P., MACE, E. M., ORANGE, J. S., CUELLAR-RODRIGUEZ, J., HICKSTEIN, D. D. & HOLLAND, S. M. (2014). GATA2 deficiency: A protean disorder of hematopoiesis, lymphatics, and immunity. *Blood*, *123*(6), 809–821. <https://doi.org/http://dx.doi.org/10.1182/blood-2013-07-515528>
- STEVEN H. SWERDLOW, ELIAS CAMPO, STEFANO A. PILERI, NANCY LEE HARRIS, HARALD STEIN REINER SIEBERT, R. A., MICHELE GHIELMINI, GILLES A. SALLES, ANDREW D. ZELENETZ, AND E. S. J., VARDIMAN, J. W., STEIN, H., STEIN, H., SWERDLOW, S. H., ... JAFFE, E. S. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*, *127*(3), 2375–2390. <https://doi.org/10.1182/blood-2016-01-643569>.The
- TASKESEN, E., BULLINGER, L., CORBACIOGLU, A., SANDERS, M. A., ERPELINCK, C. A. J., WOUTERS, B. J., ... DÖHNER, K. (2011). Prognostic impact, concurrent genetic mutations, and gene expression features of AML with CEBPA mutations in a cohort of 1182 cytogenetically normal AML patients: Further evidence for CEBPA double mutant AML as a distinctive disease entity. *Blood*, *117*(8), 2469–2475. <https://doi.org/10.1182/blood-2010-09-307280>
- TAWANA, K., DRAZER, M. W., & CHURPEK, J. E. (2018). Universal genetic testing for inherited susceptibility in children and adults with myelodysplastic syndrome and acute myeloid leukemia: Are we there yet? *Leukemia*, *32*(7), 1482–1492. <https://doi.org/10.1038/s41375-018-0051-y>
- VARDIMAN, J. W., THIELE, J., ARBER, D. A., BRUNNING, R. D., BOROWITZ, M. J., PORWIT, A., ... BLOOMFIELD, C. D. (2009). The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: Rationale and important changes. *Blood*, *114*(5), 937–951. <https://doi.org/10.1182/blood-2009-03-209262>
- WEST, A. H., GODLEY, L. A., & CHURPEK, J. E. (2014). Familial myelodysplastic syndrome/acute leukemia syndromes: A review and utility for translational investigations. *Annals of the New York Academy of Sciences*, *1310*(1), 111–118. <https://doi.org/10.1111/nyas.12346>
- WLODARSKI, M. W., HIRABAYASHI, S., PASTOR, V., STARY, J., HASLE, H., MASETTI, R., ... NIEMEYER, C. M. (2016). Prevalence, clinical characteristics, and prognosis of GATA2-related myelodysplastic syndromes in children and adolescents. *Blood*, *127*(11), 1387–1397. <https://doi.org/10.1182/blood-2015-09-669937>
- XIAO, H., SHI, J., LUO, Y., TAN, Y., HE, J., XIE, W., ZHANG, L., WANG, Y., LIU, L., WU, K., YU, X., CAI, Z., LIN, M., YE, X. & HUANG, H. (2010). Multiple mutations of CEBPA contribute to leukemogenesis: The donor origin of leukemia relapse after allogeneic hematopoietic cell transplantation. *Blood*, *116*(21), 5257–5261. <https://doi.org/10.1182/blood-2010-12-326322>.The
- ZHANG, M. Y., CHURPEK, J. E., KEEL, S. B., WALSH, T., LEE, M. K., LOEB, K. R., ... SHIMAMURA, A. (2015). Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy. *Nature Genetics*, *47*(2), 180–185. <https://doi.org/10.1038/ng.3177>

8. ANEXOS

8.1. Anexo I.

Tabla Anexo I. Diseño del panel de genes dirigido a neoplasias mieloides hereditarias: genes incluidos en el panel e información sobre su localización, la condición patológica asociada y su NM. Información extraída de las bases de datos Gene, ClinVar y PubMed, del NCBI. Chr: *chromosome*, del inglés, cromosoma. ND: no disponible. NM: identificador de la secuencia de referencia.

GENES INCLUIDOS EN EL PANEL DE GENES DIRIGIDO			
Gen	RefSeq	Localización	Condición Patológica
ABCB7	NM_004299.5	Chr X	<i>Anemia, sideroblastic, with ataxia</i>
ACD	NM_001082486.1	Chr 16	<i>Inherited bone marrow failure associated with germline mutation of ACD, the gene encoding telomere protein TPP1</i>
ADA	NM_000022.3	Chr 20	<i>Severe combined immunodeficiency due to adenosine deaminase deficiency</i>
AK2	NM_001625.3	Chr 1	<i>Human adenylate kinase 2 deficiency causes a profound hematopoietic defect associated with sensorineural deafness.</i>
AK7	ND	ND	ND
ALAS2	NM_000032.4	Chr X	<i>Anemia, sideroblastic, 1</i>
ANKRD26	NM_014915.2	Chr 10	<i>Thrombocytopenia 2</i>
AP3B1	NM_003664.4	Chr 5	<i>Hermansky-Pudlak syndrome 2; neutropenia gene</i>
ASXL1	NM_015338	Chr 20	<i>Myelodysplastic syndrome, somatic</i>
ATG2B	NM_018036.6	Chr 14	<i>A germline duplication of a region that includes this gene is associated with predisposition to myeloid malignancies; {Myeloproliferative neoplasms, familial, susceptibility to}</i>
ATM	NM_000051.3	Chr 11	<i>Lymphoma, B-cell non-Hodgkin, somatic; Lymphoma, mantle cell, somatic; T-cell prolymphocytic leukemia, somatic</i>
ATR	NM_001184	Chr 3	<i>Cutaneous telangiectasia and cancer syndrome, familial</i>
ATRX	NM_000489.4	Chr X	<i>Alpha-thalassemia/mental retardation syndrome</i>
BCOR	NM_001123383	Chr X	<i>BCOR and BCORL1 mutations in myelodysplastic syndromes (MDS). Correpresor BCL6</i>
BCORL1	NM_001184772 , NM_021946	Chr X	<i>BCOR and BCORL1 mutations in myelodysplastic syndromes (MDS). Correpresor BCL7</i>
BDKRB1	ND	Chr 14q32.2	ND
BDKRB2	ND	Chr 14q32.2	ND
BLM	NM_000057.3	Chr 15	<i>Bloom syndrome</i>
BPGM	NM_001724.4	Chr 7	<i>Erythrocytosis due to bisphosphoglycerate mutase deficiency</i>

BRCA1 (FANCS)	NM_007294.3	Chr 17	<i>Fanconi anemia, complementation group S</i>
BRIP1 (FANCI)	NM_032043.2	Chr 17	<i>Fanconi anemia, complementation group J</i>
C15ORF4 1	NM_001130010.2	Chr 15	<i>Congenital dyserythropoietic anemia type Ib</i>
CALR	NM_004343	Chr 19	<i>Myelofibrosis, somatic; Thrombocythemia, somatic</i>
CBL	NM_005188.3	Chr 11	<i>Juvenile myelomonocytic leukemia; Noonan syndrome-like disorder with or without juvenile myelomonocytic leukemia</i>
CDAN1	NM_138477.3	Chr 15	<i>Congenital dyserythropoietic anemias.</i>
CDC25C	NM_001790.4	Chr 5	ND
CEBPA	NM_004364.4	Chr 19	<i>Familial Acute Myeloid Leukemia With Mutated CEBPA</i>
CHEK2 (CHK2)	NM_007194.3	Chr 22	<i>Loss of function germline mutation; Ubiquitous expression in bone marrow</i>
CSF3R	NM_156039.3	Chr 1	<i>Neutropenia, severe congenital, 7, autosomal recessive</i>
CSNK1A1	NM_001025105.2	Chr 5	<i>differentially expressed in the bone marrow microenvironment of osteoporotic patients, providing new ideas for finding therapeutic targets for osteoporosis.</i>
CTC1	NM_025099.5	Chr 17	<i>OMIM: Cerebroretinal microangiopathy with calcifications and cysts; Pubmed: Dyskeratosis congenita (DC)</i>
CUX1	NM_181552.3	Chr 3	<i>Pubmed involved in regulation of dendritogenesis and cortical synapse formation in layer II to IV cortical neurons.</i>
CXCR4	NM_003467.2	Chr 2	<i>Myelokathexis, isolated; WHIM syndrome</i>
DDX41	NM_016222.3	Chr 5	<i>{Myeloproliferative/lymphoproliferative neoplasms, familial (multiple types), susceptibility to}</i>
DKC1	NM_001363.4	Chr X	<i>Dyskeratosis congenita, X-linked</i>
DNAJC21	NM_194283.3	Chr 5	<i>Bone Marrow Failure Syndrome 3</i>
DNMT3A	NM_175629	Chr 2	<i>Acute myeloid leukemia, somatic (tambien Tatton-Brown-Rahman syndrome)</i>
EGLN1 (PHD2)	NM_022051.2	Chr 1	<i>Erythrocytosis, familial</i>
ELANE	NM_001972.3	Chr 19	<i>Neutropenia, cyclic; Neutropenia, severe congenital 1, autosomal dominant</i>
EPAS1	NM_001430	Chr 2	<i>Mutations in this gene are associated with erythrocytosis familial type 4.</i>
EPCAM	NM_002354	Chr 2	<i>Colorectal cancer, hereditary nonpolyposis, type 8; Diarrhea 5, with tufting enteropathy, congenital</i>
EPO	NM_000799	Chr 7	<i>Diamond-Blackfan anemia-like; Erythrocytosis, familial, 5</i>
EPOR	NM_000121	Chr 19	<i>[Erythrocytosis, familial, 1]</i>
ETNK1	NM_018638	Chr 12	<i>Atypical chronic myeloid leukemia: Common cytogenetic features and mutations include</i>

			<i>trisomy 8, and mutations in SETBP1 and ETNK1</i>
ETV6	NM_001987.4	Chr 12	<i>Leukemia, acute myeloid, somatic; Thrombocytopenia 5</i>
EZH2	NM_004456	Chr 7	<i>Weaver syndrome. Remodelador de histonas (histona metil transferasa)</i>
FANCA	NM_000135.3	Chr 16	<i>Fanconi anemia, complementation group A</i>
FANCB	NM_152633.3	Chr X	<i>Fanconi anemia, complementation group B</i>
FANCC	NM_000136.2	Chr 9	<i>Fanconi anemia, complementation group C</i>
FANCD1/ BRCA2	NM_000059.3	Chr 13	<i>Fanconi anemia, complementation group D1</i>
FANCD2	NM_033084.4	Chr 3	<i>Fanconi anemia, complementation group D2</i>
FANCE	NM_021922.2	Chr 6	<i>Fanconi anemia, complementation group E</i>
FANCF	NM_022725.3	Chr 11	<i>Fanconi anemia, complementation group F</i>
FANCG	NM_004629.1	Chr 9	<i>Fanconi anemia, complementation group G</i>
FANCI	NM_001113378.1	Chr 15	<i>Fanconi anemia, complementation group I</i>
FANCL	NM_001114636.1	Chr 2	<i>Fanconi anemia, complementation group L</i>
FANCM	NM_020937.3	Chr 14	<i>A tumor suppressive DNA translocase; Premature ovarian failure 15; Spermatogenic failure 28</i>
FANCN/P ALB2	NM_024675.3	Chr 16	<i>Fanconi anemia, complementation group N</i>
FANCO/R AD51C	NM_058216.2	Chr 17	<i>Fanconi anemia, complementation group O</i>
FANCP/SL X4	NM_032444.3	Chr 16	<i>Fanconi anemia, complementation group P</i>
FANQ/E RCC4	NM_005236.2	Chr 16	<i>Fanconi anemia, complementation group Q</i>
FLT3	NM_004119	Chr 13	<i>Leukemia, acute lymphoblastic, somatic; Leukemia, acute myeloid, reduced survival in, somatic; Leukemia, acute myeloid, somatic</i>
G6PC3	NM_138387.3	Chr 17	<i>Dursun syndrome; Neutropenia, severe congenital 4, autosomal recessive</i>
GATA1	NM_002049.3	Chr X	<i>Anemia, X-linked, with/without neutropenia and/or platelet abnormalities; Leukemia, megakaryoblastic, with or without Down syndrome, somatic; Thrombocytopenia with beta-thalassemia, X-linked; Thrombocytopenia, X-linked, with or without dyserythropoietic anemia</i>
GATA2	NM_032638.4	Chr 3	<i>"Emberger syndrome; Immunodeficiency 21, {Leukemia, acute myeloid, susceptibility to}; {Myelodysplastic syndrome, susceptibility to}</i>
GFI1	NM_001127216.2	Chr 1	<i>Neutropenia, nonimmune chronic idiopathic, of adults; Neutropenia, severe congenital 2, autosomal dominant</i>
GSKIP	NM_001271904.1	Chr 14	<i>Germline duplication of ATG2B and GSKIP predisposes to familial myeloid malignancies.</i>
HAX1	NM_006118.3	Chr 1	<i>Neutropenia, severe congenital 3, autosomal recessive</i>
EPAS1	NM_001430.4	Chr 2	<i>Erythrocytosis familial type 4</i>

HSPA9	NM_004134.6	Chr 5	<i>Congenital sideroblastic anemia</i>
IDH1	NM_00128238	Chr 2	<i>{Glioma, susceptibility to, somatic}</i>
IDH2	NM_002168	Chr 15	<i>D-2-hydroxyglutaric aciduria 2</i>
IKZF1	NM_006060.5	Chr 7	<i>Immunodeficiency, common variable, 13</i>
IL2RG	NM_000206.2	Chr X	<i>X-Linked Severe Combined Immunodeficiency</i>
JAGN1	NM_032492.3	Chr 3	<i>Severe Congenital Neutropenia</i>
JAK2	NM_001322194.1	Chr 9	<i>Primary Myelofibrosis</i>
KIF23	NM_138555.3	Chr 15	<i>Congenital Dyserythropoietic Anemia; Broad expression in bone marrow</i>
KIT	NM_000222	Chr 4	<i>Leukemia, acute myeloid</i>
KLF1	NM_006563	Chr 19	<i>"Dyserythropoietic anemia, congenital, type IV; Blood group--Lutheran inhibitor; [Hereditary persistence of fetal hemoglobin]</i>
KMT2A	NM_001197104	Chr 11	<i>"Leukemia, myeloid/lymphoid or mixed-lineage; Wiedemann-Steiner syndrome</i>
KRAS	NM_033360	Chr 12	<i>Leukemia, acute myeloid</i>
LAMTOR2	NM_014017.3	Chr 1	<i>Immunodeficiency due to defect in MAPBP-interacting protein</i>
LIG4	NM_001352598.1	Chr 13	<i>LIG4 syndrome; Severe combined immunodeficiency with sensitivity to ionizing radiation</i>
LYST	NM_000081.3	Chr 1	<i>Chediak-Higashi syndrome</i>
MAD2L2	NM_001127325.1	Chr 1	<i>Fanconi anemia; Broad expression in bone marrow</i>
MBD4	NM_003925	Chr 3	<i>Genomic profiling of 3 early-onset acute myeloid leukemias (AMLs) identified germ line loss of MBD4 as an initiator of 5mC-dependent hypermutation.</i>
MCFD2	NM_001171506	Chr 2	<i>Factor V and factor VIII, combined deficiency of</i>
MECOM	NM_001105077	Chr 3	<i>Radioulnar synostosis with amegakaryocytic thrombocytopenia 2</i>
MLH1	NM_000249.3	Chr 3	<i>"Mismatch repair cancer syndrome; Colorectal cancer, hereditary nonpolyposis, type 2; Muir-Torre syndrome</i>
MPL	NM_005373.2	Chr 1	<i>Myelofibrosis with myeloid metaplasia, somatic; Thrombocythemia 2; Thrombocytopenia, congenital amegakaryocytic</i>
MRE11A	NM_005591.3	Chr 11	<i>Ataxia-telangiectasia-like disorder 1</i>
MSH2	NM_000251.2	Chr 2	<i>Colorectal cancer, hereditary nonpolyposis, type 1; Mismatch repair cancer syndrome; Muir-Torre syndrome</i>
MSH6	NM_000179.2	Chr 2	<i>"Colorectal cancer, hereditary nonpolyposis, type 5, Endometrial cancer, familial; Mismatch repair cancer syndrome</i>
NAF1	NM_138386.2	Chr 4	<i>Multiple splicing variants of Naf1/ABIN-1 transcripts and their alterations in hematopoietic tumors.</i>
NBN (NBS1)	NM_002485.4	Chr 8	<i>"Aplastic anemia; Leukemia, acute lymphoblastic; Nijmegen breakage syndrome</i>

NCK1	NM_001291999	Chr 3	<i>Bcr-Abl is the transforming principle underlying chronic myelogenous leukaemia (CML). Here, we use a functional interaction proteomics approach to map pathways by which Bcr-Abl regulates defined cellular processes. The results show that Bcr-Abl regulates the actin cytoskeleton and non-apoptotic membrane blebbing via a GADS/Slp-76/Nck1 adaptor protein pathway.</i>
NF1	NM_001042492.2	Chr 17	<i>Neurofibromatosis, type 1; Juvenile myelomonocytic leukemia (JMML) is a unique clonal hematopoietic disorder of early childhood. It is classified as an overlap myeloproliferative/myelodysplastic neoplasm by the World Health Organization (WHO) and shares some features with chronic myelomonocytic leukemia in adults. JMML pathobiology is characterized by constitutive activation of the Ras signal transduction pathway. About 90% of patients harbor molecular alterations in one of five genes (PTPN11, NRAS, KRAS, NF1 or CBL) which define genetically and clinically distinct subtypes.</i>
NHP2	NM_017838.3	Chr 5	<i>Dyskeratosis congenita; Broad expression in bone marrow</i>
NOP10	NM_018648.3	Chr 15	<i>Dyskeratosis congenita, autosomal recessive 1</i>
NPM1	"NM_001355007, NM_002520 "	Chr 5	<i>Leukemia, acute myeloid, somatic; Coexisting and cooperating mutations in NPM1-mutated acute myeloid leukemia.</i>
NRAS	NM_002524	Chr 1	<i>RAS-associated autoimmune lymphoproliferative syndrome type IV, somatic</i>
PARN	NM_002582.3	Chr 16	<i>"Dyskeratosis congenita, autosomal recessive 6; Pulmonary fibrosis and/or bone marrow failure, telomere-related, 4</i>
PAX5	NM_016734.2	Chr 9	<i>{Leukemia, acute lymphoblastic, susceptibility to, 3}</i>
PHF6	"NM_001015877, NM_032458 "	Chr X	<i>Borjeson-Forssman-Lehmann syndrome</i>
PMS2	NM_000535.5	Chr 7	<i>"Mismatch repair cancer syndrome; Colorectal cancer, hereditary nonpolyposis, type 4</i>
POT1	NM_015450.2	Chr 7	<i>"PROTECTION OF TELOMERES 1; Germline variants in the POT1 gene have recently been shown to be associated with risk of developing tumors in different tissues such as familial chronic lymphocytic leukemia; {Glioma susceptibility 9}; {Melanoma, cutaneous malignant, susceptibility to, 10}</i>

PPM1D	NM_003620	Chr 17	<i>"development of clonal hematopoiesis; in the setting of chemotherapy treatment of solid tumors, hematopoietic mutations in TP53 and PPM1D appear to contribute to outgrowth of clones that may lead to subsequent malignancy.; Breast cancer, somatic; Intellectual developmental disorder with gastrointestinal difficulties and high pain threshold</i>
PTPN11	NM_002834.4	Chr 12	<i>Leukemia, juvenile myelomonocytic, somatic</i>
PUS1	NM_025215.5	Chr 12	<i>Myopathy, lactic acidosis, and sideroblastic anemia</i>
RAB27A	NM_183236.2	Chr 15	<i>disruption of exosome secretion in AML cells through targeting Rab27a, an important regulator involved in exosome release, significantly delayed leukemia development.; Griscelli syndrome, type 2</i>
RAC2	NM_002872.4	Chr 22	<i>Neutrophil immunodeficiency syndrome</i>
RAD21	NM_006265	Chr 8	<i>"Cohesin RAD21 Gene Promoter Methylation in Patients with Chronic Lymphocytic Leukemia.; early all NPM1-mutated AML patients showed concurrent mutations in genes involved in regulation of DNA methylation (DNMT3A, TET2, IDH1, IDH2), RNA splicing (SRSF2, SF3B1), or in the cohesin complex (RAD21, SMC1A, SMC3, STAG2); Mungan syndrome; Cornelia de Lange syndrome 4</i>
RAD50	NM_005732.3	Chr 5	<i>A-AML was characterized by increased genomic complexity based on exonic variants (an average of 26 somatic mutations per sample vs 15 for E-AML). The integration of exome, copy number, and gene expression data revealed alterations in genes involved in DNA repair (eg, SLX4IP, RINT1, HINT1, and ATR) and the cell cycle (eg, MCM2, MCM4, MCM5, MCM7, MCM8, MCM10, UBE2C, USP37, CK2, CK3, CK4, BUB1B, NUSAP1, and E2F) in A-AML, which was associated with a 3-gene signature defined by PLK1 and CDC20 upregulation and RAD50 downregulation and with structural or functional silencing of the p53 transcriptional program.; Nijmegen breakage syndrome-like disorder</i>
RAD51 (FANCR)	NM_133487.3	Chr 15	<i>Fanconi anemia, complementation group R</i>
RBBP6	NM_006910.4	Chr 16	<i>Growth factor independence 1B (Gfi-1B), a transcription factor essential for the development of hematopoietic cells and differentiation of erythroid and megakaryocytic lineages; transcripts within</i>

			<i>components of the signalling cascade of immune cells (PLD1, LAMP1, HSP90, IL6ST), of the tyrosine kinase pathway (TPR, RAC3) and of the transcription factors (RAC3, CEP290, JEM-1, ATR, MYC, SMC3, RARA, RBBP6) were found to be differentially expressed in Gfi-1B overexpressing cells compared to controls.</i>
RBM8A	NM_005105.4	Chr 1	<i>Thrombocytopenia-absent radius syndrome</i>
RGS2	NM_00292	Chr 1	<i>RGS2 is an important target gene of Flt3-ITD mutations in AML and functions in myeloid differentiation and leukemic transformation.</i>
RHAG	NM_000324	Chr 6	<i>Anemia, hemolytic, Rh-null, regulator type</i>
RMRP	NR_003051.3	Chr 9	<i>Anauxetic dysplasia 1; Cartilage-hair hypoplasia; Metaphyseal dysplasia without hypotrichosis; (Extended follow-up of the Finnish cartilage-hair hypoplasia cohort confirms high incidence of non-Hodgkin lymphoma and basal cell carcinoma.)</i>
RNF168	NM_152617.3	Chr 3	<i>RIDDLE syndrome</i>
RPL10	NM_001303624.1	Chr X	<i>"The pathogenic role of the recurrent R98S mutation in ribosomal protein L10 (RPL10 R98S) found in T-cell acute lymphoblastic leukemia (T-ALL); Mental retardation, X-linked, syndromic, 35; {Autism, susceptibility to, X-linked 5}</i>
RPL11	NM_000975.4	Chr 1	<i>Diamond-Blackfan anemia 7</i>
RPL15	NM_001253383.2	Chr 3	<i>Diamond-Blackfan anemia 12</i>
RPL26	NM_000987.4	Chr 17	<i>Diamond-Blackfan anemia 11</i>
RPL27	NM_001349922	Chr 17	<i>Diamond-Blackfan anemia 16</i>
RPL35A	NM_000996.3	Chr 3	<i>Diamond-Blackfan anemia 5</i>
RPL5	NM_000969.4	Chr 1	<i>Diamond-Blackfan anemia 6</i>
RPL9	"NM_000661, NM_001024921 "	Chr 4	ND
RPS10	NM_001203245.2	Chr 6	<i>"Diamond-Blackfan anemia 9</i>
RPS14	NM_001025071.1	Chr 5	<i>5q minus (5q-) syndrome is a type of bone marrow disorder called myelodysplastic syndrome (MDS)</i>
RPS17	NM_001021.5 *ribosome	Chr 15	<i>Diamond-Blackfan anemia 4</i>
RPS19	NM_001022.3	Chr 19	<i>Diamond-Blackfan anemia 1</i>
RPS24	NM_001142285.1	Chr 10	<i>Diamond-blackfan anemia 3</i>
RPS26	NM_001029.4	Chr 12	<i>Diamond-Blackfan anemia 10</i>
RPS27	"NM_001030 , NM_001349946, NM_001349947 "	Chr1	<i>Diamond-Blackfan anemia 17</i>
RPS29	NM_001030001.3	Chr 14	<i>Diamond-Blackfan anemia 13</i>
RPS7	NM_001011.3	Chr 2	<i>Diamond-Blackfan anemia 8</i>
RTEL1	NM_032957.4	Chr 20	<i>Germline mutations of regulator of telomere elongation helicase 1, RTEL1, in Dyskeratosis congenita.</i>

RUNX1	NM_001754.4	Chr 21	<i>Leukemia, acute myeloid; Platelet disorder, familial, with associated myeloid malignancy</i>
SAMD9	NM_001193307.1	Chr 7	<i>"MIRAGE syndrome; Tumoral calcinosis, familial, normophosphatemic</i>
SAMD9L	NM_152703.4	Chr 7	<i>"Ataxia-pancytopenia syndrome</i>
SBDS	NM_016038.3	Chr 7	<i>"{Aplastic anemia, susceptibility to}; Shwachman-Diamond syndrome</i>
SBF2	NM_030962.3	Chr 11	<i>Charcot-Marie-Tooth disease, type 4B2; The data presented indicate that B cells from patients with chronic lymphocytic leukemia (CLL) may produce suppressive factors capable of inhibiting T and B cell proliferation and Ig secretion. Two factors with the molecular weight 65-75 KDa (SBF1) and 15-25 KDa (SBF2) have been characterized. SBFs are primarily produced by B cells with a low flow density. There is a reverse dependence between the capacity of leukemic B cells for spontaneous Ig secretion and their ability for the production of suppressive factors.</i>
SETBP1	NM_015559	Chr 18	<i>"Mental retardation, autosomal dominant 29; Schinzel-Giedion midface retraction syndrome</i>
SF3B1	NM_012433	Chr 2	<i>Myelodysplastic syndrome, somatic</i>
SH2B3	NM_005475.2	Chr 12	<i>"Myelofibrosis, somatic; Thrombocythemia, somatic; Erythrocytosis, somatic</i>
SMC1A	NM_006306	Chr X	<i>"Hematologic;; Cornelia de Lange syndrome 2</i>
SMC3	NM_005445	Chr 10	<i>"Articles on SLC25A38;; Cornelia de Lange syndrome 3</i>
SMOH	NM_005631	Chr 7	<i>"Basal cell carcinoma, somatic; Curry-Jones syndrome, somatic mosaic</i>
SRP54	ND	ND	<i>Shwachman-Diamond syndrome (SDS) is a rare inherited disease mainly caused by mutations in the Shwachman-Bodian-Diamond Syndrome (SBDS) gene. However, it has recently been reported that other genes, including DnaJ heat shock protein family (Hsp40) member C21 (DNAJC21), elongation factor-like 1 (EFL1) and signal recognition particle 54 (SRP54) are also associated with an SDS-like phenotype.</i>
SRP72	NM_006947.3	Chr 4	<i>Bone marrow failure syndrome 1</i>
SRSF2	NM_003016	Chr 17	ND
STAG2	NM_001042749	Chr X	<i>NUCLEOPHOSMIN1 (NPM1) is the most frequently mutated gene in acute myeloid leukemia. Notably, NPM1 mutations are always accompanied by additional mutations such as those in cohesin genes RAD21, SMC1A, SMC3, STAG2</i>

TAZ	NM_000116.4	Chr X	<i>Barth syndrome</i>
TCIRG1	NM_006019.3	Chr 11	<i>Osteopetrosis, autosomal recessive 1; Severe Congenital Neutropenia ??; infantile malignant osteopetrosis (IMO). Genetic mutation in T-cell, immune regulator 1 (TCIRG1) was identified, confirming the diagnosis of IMO</i>
TCL1A	NM_021966	Chr 14	<i>Leukemia/lymphoma, T-cell</i>
TERC	NR_001566.1	Chr 3	<i>Dyskeratosis congenita, autosomal dominant 1; {Aplastic anemia}; {Pulmonary fibrosis, idiopathic, susceptibility to}</i>
TERF1	NM_017489.2	Chr 8	<i>TELOMERIC REPEAT-BINDING FACTOR 1</i>
TERF2	NM_005652.4	Chr 16	<i>TELOMERIC REPEAT-BINDING FACTOR 2</i>
TERT	NM_198253.2	Chr 5	<i>"{Leukemia, acute myeloid}; {Dyskeratosis congenita, autosomal dominant 2}; {Dyskeratosis congenita, autosomal recessive 4}; {Pulmonary fibrosis and/or bone marrow failure, telomere-related, 1}; {Melanoma, cutaneous malignant, 9}</i>
TET2	NM_001127208.2	Chr 4	<i>Myelodysplastic syndrome, somatic</i>
TINF2	NM_001099274.1	Chr 14	<i>TINF2 mutations result in very short telomeres: analysis of a large cohort of patients with dyskeratosis congenita and related bone marrow failure syndromes.</i>
TP53	NM_000546.5	Chr 17	<i>Bone marrow failure syndrome 5</i>
TRNT1	NM_182916.2	Chr 3	<i>Sideroblastic anemia with B-cell immunodeficiency, periodic fevers, and developmental delay</i>
TSR2	NM_058163.2	Chr X	<i>Diamond-Blackfan anemia 14 with mandibulofacial dysostosis</i>
U2AF1	NM_006758	Chr 21	<i>Clonal hematopoiesis is frequently present in otherwise healthy individuals and may persist for many years. Though largely asymptomatic, carrying these somatic mutations confers a small but significantly increased risk of leukemic transformation, affecting 0.5-1% carriers per year; although most genes confer an increased risk of transformation, mutations in TP53 and U2AF1 appear to carry a particularly high risk for transformation.</i>
UBE2T (FANCT)	NM_014176.3	Chr 1	<i>Fanconi anemia, complementation group T</i>
USB1 (C16orf57)	NM_024598.3	Chr 16	<i>Poikiloderma with neutropenia</i>
VHL	NM_000551.3	Chr 3	<i>Familial Erythrocytosis</i>
VPS13B	NM_017890.4	Chr 8	<i>Cohen syndrome;; Rearrangement of VPS13B, a causative gene of Cohen syndrome, in a case of RUNX1-RUNX1T1 leukemia with t(8;12;21).</i>

VPS45	NM_007259.4	Chr 1	<i>Neutropenia, severe congenital, 5, autosomal recessive</i>
WAS	NM_000377.2	Chr X	<i>"Neutropenia, severe congenital, X-linked; Thrombocytopenia, X-linked; Thrombocytopenia, X-linked, intermittent; Wiskott-Aldrich syndrome</i>

8.2. Anexo II.

Tabla Anexo II. Zonas de mejora del panel de genes dirigido. Se incluyen la localización cromosómica, las posiciones concretas iniciales y finales, el gen a que corresponde y el NM junto con el exón correspondiente. Chr: *chromosome*, del inglés, cromosoma. NM: identificador de la secuencia de referencia.

POSICIONES CROMOSÓMICAS MAL CUBIERTAS				
Localización	Posición inicial	Posición final	Gen	NM con exón
Chr 9	5077452	5077580	JAK2	NM_001322195.ex.14
Chr 7	6017218	6017388	PMS2	NM_001322004.ex.13
Chr 7	6018226	6018327	PMS2	NM_001322004.ex.12
Chr 7	6043320	6043423	PMS2	NM_001322004.ex.3
Chr 7	6043602	6043689	PMS2	NM_001322004.ex.2
Chr 17	8151321	8151413	CTC1	NM_025099.ex.1
Chr 3	10085167	10085276	FANCD2	NM_001319984.ex.13
Chr 3	10085512	10085548	FANCD2	NM_001319984.ex.14
Chr 3	10091057	10091189	FANCD2	NM_001319984.ex.17
Chr 3	10105475	10105595	FANCD2	NM_001319984.ex.21
Chr 3	10106039	10106113	FANCD2	NM_001319984.ex.22
Chr 3	10106412	10106559	FANCD2	NM_001319984.ex.23
Chr 3	10107547	10107663	FANCD2	NM_001319984.ex.25
Chr 16	14678605	14678648	PARN	NM_001242992.ex.14
Chr 16	14693760	14693817	PARN	NM_001242992.ex.11
Chr 16	14702137	14702176	PARN	NM_001242992.ex.8
Chr X	14871160	14871289	FANCB	NM_001018113.ex.6
Chr X	14875983	14876076	FANCB	NM_001018113.ex.5
Chr X	15833799	15834013	ZRSR2	NM_005089.ex.8
Chr 16	24557483	24557583	RBBP6	NM_032626.ex.2
Chr 10	27311486	27311616	ANKRD26	NM_014915.ex.29
Chr 10	27313375	27313488	ANKRD26	NM_014915.ex.28
Chr 10	27333085	27333119	ANKRD26	NM_014915.ex.19
Chr 10	27342248	27342319	ANKRD26	NM_014915.ex.16
Chr 10	27371731	27371762	ANKRD26	NM_014915.ex.6
Chr 22	29083730	29083974	CHEK2	NM_007194.ex.15

Chr 22	29085122	29085203	CHEK2	NM_007194.ex.14
Chr 22	29090019	29090105	CHEK2	NM_007194.ex.13
Chr 22	29091114	29091230	CHEK2	NM_007194.ex.12
Chr 22	29091697	29091861	CHEK2	NM_007194.ex.11
Chr 22	29105993	29106047	CHEK2	NM_007194.ex.7
Chr 22	29115382	29115473	CHEK2	NM_007194.ex.5
Chr 17	29541468	29541603	NF1	NM_001128147.ex.13
Chr 17	29548867	29549782	NF1	NM_001128147.ex.15
Chr 17	29557859	29557943	NF1	NM_001042492.ex.24
Chr 17	29579955	29580018	NF1	NM_001042492.ex.31
Chr 8	30925773	30925843	WRN	NM_000553.ex.7
Chr 8	30941214	30941295	WRN	NM_000553.ex.10
Chr 8	30942681	30942762	WRN	NM_000553.ex.11
Chr 8	30947980	30948048	WRN	NM_000553.ex.14
Chr 8	31000141	31000217	WRN	NM_000553.ex.27
Chr 13	32903579	32903629	BRCA2	NM_000059.ex.8
Chr 13	32918694	32918790	BRCA2	NM_000059.ex.12
Chr 13	32920963	32921033	BRCA2	NM_000059.ex.13
Chr 21	36265221	36265260	RUNX1	NM_001754.ex.3
Chr 3	37036320	37036444	MLH1	NM_001354619.ex.2
Chr 3	37083758	37083822	MLH1	NM_001354629.ex.14
Chr 15	40994003	40994124	RAD51	NM_133487.ex.4
Chr 17	41249260	41249306	BRCA1	NM_007298.ex.7
Chr 14	45624575	45624662	FANCM	NM_001308134.ex.8
Chr 14	45650632	45650727	FANCM	NM_001308133.ex.14
Chr 14	45650839	45650908	FANCM	NM_001308133.ex.15
Chr 14	45654419	45654576	FANCM	NM_001308133.ex.17
Chr 14	45656983	45657090	FANCM	NM_001308133.ex.18
Chr 2	47602372	47602438	EPCAM	NM_002354.ex.4
Chr 14	50050289	50050393	RPS29	NM_001032.ex.3
Chr 14	50050290	50050393	RPS29	NM_001351375.ex.3
Chr 4	57352490	57352563	SRP72	NM_006947.ex.11
Chr 4	57361522	57361560	SRP72	NM_006947.ex.17
Chr 7	66456123	66456288	SBDS	NM_016038.ex.4
Chr 7	66458203	66458404	SBDS	NM_016038.ex.3
Chr 7	66459198	66459328	SBDS	NM_016038.ex.2
Chr 8	73926129	73926225	TERF1	NM_017489.ex.2
Chr 8	73937056	73937206	TERF1	NM_017489.ex.5
Chr 8	73951350	73951454	TERF1	NM_017489.ex.9
Chr X	76872080	76872198	ATRX	NM_138270.ex.21
Chr X	76889053	76889200	ATRX	NM_138270.ex.17
Chr X	76907603	76907843	ATRX	NM_138270.ex.14
Chr X	76912049	76912143	ATRX	NM_138270.ex.12

Chr X	76940430	76940498	ATRX	NM_138270.ex.7
Chr X	76952064	76952192	ATRX	NM_138270.ex.5
Chr X	76954061	76954117	ATRX	NM_138270.ex.3
Chr 5	77396776	77396849	AP3B1	NM_001271769.ex.21
Chr 5	77461433	77461496	AP3B1	NM_001271769.ex.12
Chr 5	77521365	77521432	AP3B1	NM_001271769.ex.6
Chr 15	82821157	82821289	RPS17	NR_111943.ex.4
Chr 15	82822713	82822779	RPS17	NR_111943.ex.3
Chr 15	82823287	82823393	RPS17	NR_111943.ex.2
Chr 15	82824388	82824865	RPS17	NR_111943.ex.1
Chr 15	83205500	83205632	RPS17	NR_111943.ex.4
Chr 15	83207056	83207122	RPS17	NR_111943.ex.3
Chr 15	83207630	83207736	RPS17	NR_111943.ex.2
Chr 15	83208731	83209208	RPS17	NR_111943.ex.1
Chr 15	89816607	89816700	FANCI	NM_001113378.ex.11
Chr 8	90949253	90949303	NBN	NM_002485.ex.15
Chr 8	90960051	90960120	NBN	NM_002485.ex.12
Chr 15	91310139	91310253	BLM	NM_001287246.ex.11
Chr 14	96761818	96761897	ATG2B	NM_018036.ex.35
Chr 14	96777878	96777955	ATG2B	NM_018036.ex.27
Chr 14	96779665	96779772	ATG2B	NM_018036.ex.24
Chr 14	96810990	96811093	ATG2B	NM_018036.ex.4
Chr 9	98002930	98003025	FANCC	NM_001243743.ex.4
Chr 8	100146859	100146955	VPS13B	NM_017890.ex.9
Chr 15	83207630	83207736	RPS17	NR_111943.ex.2
Chr 15	83208731	83209208	RPS17	NR_111943.ex.1
Chr 15	89816607	89816700	FANCI	NM_001113378.ex.11
Chr 8	90949253	90949303	NBN	NM_002485.ex.15
Chr 8	90960051	90960120	NBN	NM_002485.ex.12
Chr 15	91310139	91310253	BLM	NM_001287246.ex.11
Chr 14	96761818	96761897	ATG2B	NM_018036.ex.35
Chr 14	96777878	96777955	ATG2B	NM_018036.ex.27
Chr 14	96779665	96779772	ATG2B	NM_018036.ex.24
Chr 14	96810990	96811093	ATG2B	NM_018036.ex.4
Chr 9	98002930	98003025	FANCC	NM_001243743.ex.4
Chr 8	100146859	100146955	VPS13B	NM_017890.ex.9
Chr 11	108159703	108159830	ATM	NM_001351834.ex.29
Chr 11	108204612	108204695	ATM	NM_001351834.ex.55
Chr 10	112361723	112361936	SMC3	NM_005445.ex.25
Chr X	123156380	123156521	STAG2	NM_001042750.ex.3
Chr X	123159689	123159768	STAG2	NM_001042750.ex.4
Chr X	123176418	123176495	STAG2	NM_001042750.ex.7
Chr X	123184970	123185069	STAG2	NM_001042750.ex.12

Chr X	123199725	123199796	STAG2	NM_001042750.ex.21
Chr X	123200024	123200112	STAG2	NM_001042750.ex.22
Chr X	123202413	123202506	STAG2	NM_001042750.ex.24
Chr X	123211806	123211908	STAG2	NM_001042750.ex.27
Chr 7	124537218	124537266	POT1	NR_003102.ex.5
Chr 5	131894975	131895059	RAD50	NM_005732.ex.2
Chr 5	131944306	131944417	RAD50	NM_005732.ex.17
Chr 5	131944808	131944901	RAD50	NM_005732.ex.18
Chr X	133512034	133512136	PHF6	NM_001015877.ex.3
Chr X	133527938	133527982	PHF6	NM_001015877.ex.5
Chr X	133549045	133549150	PHF6	NM_001015877.ex.8
Chr 3	136649316	136649436	NCK1	NM_001190796.ex.1
Chr 1	145510727	145510938	GNRHR2	NR_104033.ex.3
Chr 1	150115015	150115109	VPS45	NM_001279353.ex.13
Chr 7	152357785	152357867	XRCC2	NM_005431.ex.2
Chr 4	164066933	164067016	NAF1	NM_138386.ex.4
Chr 3	168825713	168825740	MECOM	NM_004991.ex.10
Chr 2	198283520	198285266	SF3B1	NM_001308824.ex.4
Chr 1	235875356	235875497	LYST	NM_001301365.ex.43
Chr 1	235963619	235963686	LYST	NM_001301365.ex.10