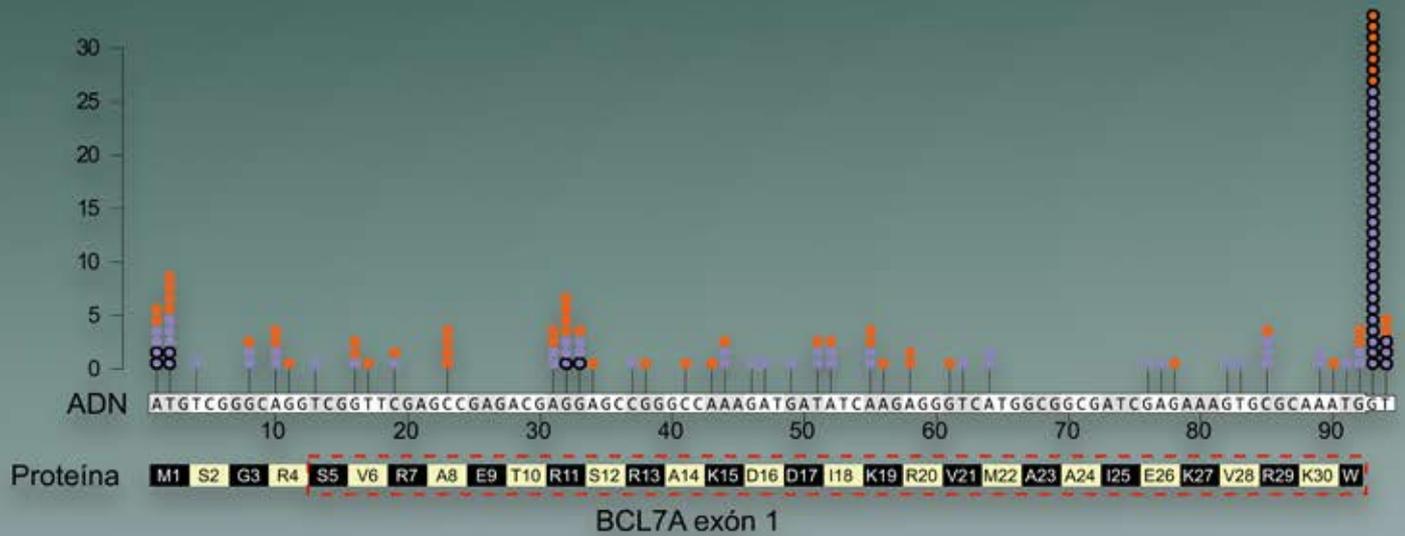


TESIS DOCTORAL

ESTUDIO DE LAS IMPLICACIONES DE BCL7A EN EL DESARROLLO TUMORAL



Carlos Baliñas Gavira



UNIVERSIDAD
DE GRANADA



FROM UNIVERSIDAD DE GRANADA - IUTA DE ANDALUCÍA
CENTRE FOR GENOMICS AND ONCOLOGICAL RESEARCH



**UNIVERSIDAD
DE GRANADA**

DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I

ESCUELA DE DOCTORADO DE CIENCIAS DE LA SALUD

**PROGRAMA DE DOCTORADO EN BIOQUÍMICA Y
BIOLOGÍA MOLECULAR (B16.56.1)**

TESIS DOCTORAL

**ESTUDIO DE LAS IMPLICACIONES DE BCL7A
EN EL DESARROLLO TUMORAL**

Carlos Baliñas Gavira

Director de Tesis:

Pedro P. Medina Vico

Granada, a 18 de diciembre de 2020



CENTRO PFIZER-UNIVERSIDAD DE GRANADA-JUNTA DE ANDALUCÍA
DE **GENÓMICA E INVESTIGACIÓN ONCOLÓGICA**

Gene Expression Regulation and Cancer Group (CTS-993)

Editor: Universidad de Granada. Tesis Doctorales
Autor: Carlos Baliñas Gavira
ISBN: 978-84-1306-739-1
URI: <http://hdl.handle.net/10481/65386>

AGRADECIMIENTOS

Y aquí estoy el último día echando la vista atrás, intentando recordar todo lo que me ha llevado hasta aquí. Son muchos momentos, situaciones, detalles... pero, sobre todo, muchas personas que se han cruzado e incluso mantenido en el camino.

Y una vez te pones a mirar atrás, no sabes muy bien donde parar. Es fácil acordarse de los más lejanos, quienes fueron los primeros en plantar la semilla de la curiosidad científica en mí. Me refiero a mis profesores del colegio, en especial a mi profesor de Química D. Mario que con sus “¡Baliñas, estuuuuuuudia! buscaba sacar lo mejor de mí. También, a mi querida profesora de Biología D^a Leocricia, quien en sus clases conseguía que mi mente dispersa se quedase obnubilada con la cantidad de cosas que una simple célula podía hacer.

Del colegio llegué a la Complu, lugar que siguió siendo un patio de recreo y, al mismo tiempo, nos fue curtiendo en el manejo de la espátula, ahora convertida en pipeta. Primero los químicos, luego los bioquímicos, los erasmus y finalmente los inmunólogos. No hay espacio para nombrar a cada uno de vosotros, pero quiero agradecerlos a todos el haber formado parte de esta gran etapa. Y aunque la vida nos aleja, siempre quedarán los reencuentros para ponernos al día.

También queda lejos el momento en el que hice las maletas para irme a Granada. Gracias a Pedro por darme la oportunidad y confiar en mí. Gracias al LAB21, conocidos como los “PMVs”. Hemos sido muchos arrimando el hombro para que todo vaya hacia adelante. Gracias Marta por escucharme, gracias Maribel por pelear los experimentos a mi lado, gracias Paola por introducirme al yoga en mis momentos más oscuros y compartir tu optimismo. Todas vosotras lleváis conmigo casi desde el principio y habéis sido grandes compañeras. Luego vinieron las nuevas generaciones de becarios: Alberto, Dani, Juansan, JuanRo... habéis hecho del laboratorio un lugar donde los días largos se hacen más fáciles. Gracias también a Carlos “el gallego” por su sinceridad, sencillez y apoyo.

También han sido muchos los estudiantes que han “disfrutado” de mi compañía para realizar sus TFGs o TFM: Cristina, Anabel, Ana, Carmen, Juanma, Thomas, Pablo... Espero que tengáis un buen recuerdo de la experiencia. Para mí ha sido muy enriquecedor y os agradezco vuestra ayuda y dedicación. También quería agradecer a la gente del CNIC, a Almudena, Virginia y Ángel, que me hicieron un hueco en su laboratorio para ponerme al día sobre las células B. Y también agradecer enormemente a José Ángel Climent todas las molestias que se tomó en escucharme y guiarme cuando la ciencia me estaba dando una bofetada.

Gracias a Álvaro. Me siento muy afortunado de haberte tenido a mi lado durante estos últimos años. Has aguantado chapas muy intensas y has buscado solución a muchos de mis problemas. Gracias por tu dedicación infinita, vales mucho.

Gracias al LAB21 he podido conocer a dos personas muy especiales. Gracias Joel porque has sido mucho más que un compañero de trabajo. Tu alegría desbordante es el ingrediente que todo laboratorio necesita. Gracias Santi por ser un apoyo durante todos estos años, por enseñarme que la ciencia es pasión y que la pasión es la esencia de la vida.

Y más allá de las cuatro poyatas del LAB21, en Genyo he encontrado una gran familia. Muchas gracias a todos, han sido muchas risas y buenos momentos: el Corpus, los Escullos, las fiestas en la terraza, las barbacoas a lo grande o en *petit comité*, las cervecitas eternas de los viernes... No todo ha sido trabajar y no hay duda de que... para liberar la mente, el cuerpo y lo que haga falta, he tenido a unos compañeros Genyales.

La ciudad de Granada me ha dado mucho y al mismo tiempo me alejó un poco de mi gente, los del barrio, los de toda la vida... Pero no hay distancia suficiente para que la complicidad se apague. Gracias por nunca cuestionar que tenga que ir un sábado a darle de comer a las celulillas. Sé que tenéis grandes esperanzas en mí... yo solo tengo la esperanza de teneros siempre ahí, sea donde sea, para sentirme como en casa.

Gracias a mi familia, a Charo, Rocío, Nieves, José Manuel y a todos los Baliñas-Gavira et al. por celebrar mis logros y por animarme siempre a seguir.

Gracias a mi hermano, Alejandro, porque siempre has sido un gran referente para mí como ahora lo serás para Roque y Tirso. Con hermanos mayores así nunca será un problema ser un *Peck*.

Gracias a mis padres, Carmen y Antonio. Gracias por vuestro apoyo incondicional. Vuestro respaldo hace que mis metas no tengan techo. Gracias por darme siempre la libertad y el apoyo para cometer mis propios errores y por ayudarme a salir de ellos. Gracias por vuestro cariño.

Gracias a la ciencia por permitirme generar lentivirus. Sin esos títulos virales igual no hubiese podido atraer los ojitos de Julia. Gracias a Julia por ser mi compañera de un viaje que no ha hecho nada más que empezar. Gracias por darle color a mis ideas y por transmitirme calma y serenidad en los momentos duros. Que la vida continúe, mientras sea a tu lado.

A mis padres

RESUMEN

En los últimos años, el estudio del complejo remodelador de la cromatina SWI/SNF ha centrado gran interés en el campo de la oncología. Diferentes estudios de secuenciación en genomas tumorales han puesto de manifiesto la alta prevalencia de mutaciones de las distintas subunidades del complejo SWI/SNF, que alcanza una frecuencia mutacional global superior al 20% entre todos los tipos de tumores analizados. Además, se ha definido que el complejo SWI/SNF ejerce un papel supresor tumoral en determinados tumores sólidos, ya que mutaciones de pérdida de función promueven una desregulación epigenética que favorece el desarrollo tumoral.

BCL7A ha sido una de las últimas subunidades que han sido asociadas al complejo SWI/SNF, pues no presenta homólogos en los complejos SWI/SNF en levaduras. A diferencia de otras subunidades que presentan mutaciones en distintos tipos de tumores, las mutaciones en *BCL7A* parecen concentrarse en neoplasias hematológicas. Sin embargo, se desconocen las implicaciones funcionales que podrían tener dichas mutaciones. Por otro lado, a pesar de que la función biológica de *BCL7A* es una incógnita a día de hoy, los patrones de expresión génica de *BCL7A* sugieren que podría tener una función específica durante el proceso de maduración de los linfocitos B.

Esta tesis doctoral nació con el objetivo de dilucidar si *BCL7A* podría actuar como supresor tumoral en neoplasias hematológicas. Para ello, se analizó el patrón mutacional de *BCL7A* en diversas líneas celulares tumorales de origen hematológico. Este estudio inicial mostró que las mutaciones de *BCL7A* eran frecuentes en líneas celulares derivadas de pacientes con linfoma difuso de células B grandes (DLBCL). El DLBCL es el subtipo de linfoma no Hodgkin más común en la población y se caracteriza por presentar mutaciones en diferentes genes relacionados con la modificación de la cromatina y el control epigenético. Con el fin de analizar más a fondo el patrón mutacional de *BCL7A* en DLBCL, se extendió el estudio a un total de 41 líneas celulares y 38 muestras de pacientes de DLBCL.

Nuestro análisis mutacional de *BCL7A* en DLBCL mostró una prevalencia mutacional del 17% en líneas celulares y del 5,5% en muestras de pacientes. Un análisis más detallado de las mutaciones detectadas, reveló que *BCL7A* presentaba un patrón de inactivación bialélica, característica propia de genes supresores tumorales. Además, se detectó la presencia de una serie de mutaciones que se concentraban en la secuencia conservada del sitio donador de *splicing* del intrón 1, la cual no había sido tenido en cuenta previamente en otros estudios de secuenciación

masiva. Al extender el análisis mutacional a cohortes externas con un total de 1.575 pacientes de DLBCL, comprobamos que las mutaciones en el sitio donador de *splicing* suponían el 30% de todas las mutaciones de *BCL7A*.

A continuación, al estudiar las posibles implicaciones moleculares de las mutaciones en el sitio de *splicing*, descubrimos que estas mutaciones generaban un *splicing* aberrante. Debido a la activación de un sitio donador de *splicing* críptico dentro del exón 1, se generaba un ARN mensajero maduro. Este transcrito da lugar a una forma mutante de *BCL7A* que pierde 27 aminoácidos de la región amino-terminal ($\Delta 27$ -*BCL7A*).

Posteriormente, decidimos analizar las consecuencias funcionales de las mutaciones de *splicing* detectadas en *BCL7A*. Por un lado, mediante experimentos de inmunoprecipitación y espectrometría de masas, se ha caracterizado que la pérdida del dominio amino-terminal (*BCL7_Nt*) impide que *BCL7A* pueda unirse al complejo SWI/SNF y, por tanto, podría estar afectando a la función global del complejo SWI/SNF. Por otro lado, se han llevado a cabo experimentos funcionales de restauración en modelos celulares de DLBCL que carecen de la expresión de *BCL7A* salvaje debido a su inactivación bialélica. La restauración de la expresión de *BCL7A* en este contexto ejerce un fenotipo supresor tumoral tanto en estudios funcionales *in vitro* como en xenotrasplantes *in vivo*. Además, los análisis de expresión génica mediante RNA-Seq han revelado que *BCL7A* participa en vías de activación de células B y que es capaz de regular la expresión de genes relacionados con el ciclo celular como *CDKN1A* (p21).

El DLBCL es una neoplasia maligna muy heterogénea donde el 40% de los pacientes desarrollan enfermedad refractaria y terminan sucumbiendo a la enfermedad. Es por ello que el tratamiento de estos pacientes avanza hacia un escenario que requiere del desarrollo de nuevas terapias dirigidas. Nuestro estudio ha conseguido por primera vez demostrar un papel supresor tumoral de una subunidad del complejo SWI/SNF en DLBCL. Además, hemos mostrado cómo el complejo SWI/SNF se encuentra mutado en más de la mitad de los pacientes del subtipo de centro germinal (GCB)-DLBCL. Todos estos resultados refuerzan la idea de que el complejo SWI/SNF juega un papel relevante en la patogénesis del DLBCL y sirven de apoyo para nuevas aproximaciones terapéuticas que mejoren el pronóstico y tratamiento personalizado de estos pacientes mediante el uso de terapias epigenéticas que están actualmente en desarrollo.

ÍNDICE

ÍNDICE DE FIGURAS.....
ABREVIATURAS
1. INTRODUCCIÓN	1
1.1 CÁNCER	3
1.2 NEOPLASIAS HEMATOLÓGICAS	4
1.3 LINFOMA DISFUSO DE CÉLULAS B GRANDES	6
1.3.1 <i>Biología de las células B: reacción de centro germinal</i>	7
1.3.2 <i>Regulación transcripcional de la reacción de CG</i>	9
1.3.3 <i>Clasificación molecular del DLBCL</i>	10
1.3.4 <i>Tratamiento</i>	11
1.4 REMODELACIÓN DE LA CROMATINA Y COMPLEJO SWI/SNF	13
1.4.1 <i>Complejos remodeladores de la cromatina</i>	13
1.4.2 <i>Complejo SWI/SNF: Estructura y función</i>	14
1.4.3 <i>Complejo SWI/SNF y sus implicaciones en cáncer</i>	17
1.4.3 <i>Aproximaciones terapéuticas en tumores mutantes para SWI/SNF</i>	19
1.5 BCL7A	22
1.5.1 <i>BCL7A: asociación al complejo SWI/SNF e historial mutacional</i>	22
1.5.2 <i>La expresión de BCL7A parece modularse durante la reacción de CG</i>	23
2. JUSTIFICACIÓN E HIPÓTESIS	25
3. OBJETIVOS.....	29
4. MATERIALES Y MÉTODOS.....	33
4.1 MUESTRAS BIOLÓGICAS: LÍNEAS CELULARES Y PACIENTES	35
4.1.1 <i>Colección de líneas celulares para el estudio piloto</i>	35
4.1.2 <i>Panel de líneas celulares de DLBCL</i>	36
4.1.3 <i>Muestras de pacientes de DLBCL y FL</i>	36
4.2 EXTRACCIÓN DE ADN, ARN Y TRANSCRIPCIÓN INVERSA.....	36
4.3 ANÁLISIS DE EXPRESIÓN DE BCL7A EN EL ESTUDIO PILOTO.....	37
4.4 ANÁLISIS DE VARIACIÓN DE NÚMERO DE COPIAS	37
4.5 ESTUDIOS DE SECUENCIACIÓN EN LÍNEAS CELULARES Y COHORTE PROPIA DE PACIENTES DERIVADOS DE DLBCL Y FL ..	38
4.6 ESTUDIOS DE SECUENCIACIÓN EN COHORTES EXTERNAS DE PACIENTES DE DLBCL.....	39
4.6.1 <i>Descarga de datos</i>	39
4.6.2 <i>Detección y anotación de variantes</i>	40
4.7 ESTUDIO DE LA IMPLICACIÓN DE AID EN EL PATRÓN MUTACIONAL DE BCL7A	40
4.8 CULTIVOS CELULARES	42
4.9 GENERACIÓN DE VECTORES LENTIVIRALES.....	42

4.10 PRODUCCIÓN Y TITULACIÓN DE LAS PARTÍCULAS LENTIVIRALES	43
4.11 ENSAYOS DE CRECIMIENTO CELULAR COMPETITIVO	44
4.12 XENOTRASPLANTE DE LA LÍNEA CELULAR OCI-LY1 EN RATONES NSG	44
4.13 ANÁLISIS DE EXPRESIÓN GÉNICA MEDIANTE RNA-SEQ	45
4.14 ANÁLISIS DE EXPRESIÓN DE BCL7A MEDIANTE RT-QPCR	47
4.15 WESTERN BLOT	47
4.16 INMUNOPRECIPITACIÓN Y ESPECTROMETRÍA DE MASAS (LC-MS/MS)	48
4.16.1 Preparación de las muestras para el análisis proteómico	49
4.16.2 Espectrometría de masas (Q Exactive™)	49
4.16.3 Análisis de los datos generados por LC-MS/MS	50
4.17 DATOS GENERADOS EN ESTE ESTUDIO ALMACENADOS EN REPOSITORIOS	50
5. RESULTADOS	53
5.1 ESTUDIO PILOTO EN DISTINTAS NEOPLASIAS HEMATOLÓGICAS	55
5.2 ESTUDIO MUTACIONAL EN DLBCL	59
5.2.1 Variantes genéticas en BCL7A en líneas celulares	60
5.2.1.1 Mutaciones con cambio de sentido	60
5.2.1.2 Mutación puntual en el codón de inicio de la traducción	61
5.2.1.3 Mutaciones de splicing	62
5.2.1.4 Deleción en HBL1	62
5.2.2 Variantes genéticas en BCL7A en muestras de pacientes	64
5.2.3 SNPs en regiones codificantes de BCL7A	64
5.2.4 Variantes genéticas en regiones no codificantes	66
5.3 CARACTERIZACIÓN MOLECULAR DE LOS TRANSCRITOS DERIVADOS DE LAS MUTACIONES DE SPLICING	66
5.4 EXPRESIÓN DE BCL7A A NIVEL DE PROTEÍNA	71
5.5 ESTUDIO MUTACIONAL DE BCL7A EN COHORTES EXTERNAS	75
5.6 LAS MUTACIONES EN BCL7A SE CONCENTRAN EN EL EXÓN 1 Y PRESENTAN UN PATRÓN DE INACTIVACIÓN BIALÉLICA .	78
5.7 LA ACCIÓN DE LA ENZIMA AID NO JUSTIFICA LA TOTALIDAD DE MUTACIONES QUE PRESENTA BCL7A	84
5.8 EL DOMINIO AMINO-TERMINAL DE BCL7A ES CRÍTICO PARA SU INTERACCIÓN CON EL COMPLEJO SWI/SNF	87
5.9 LA RESTAURACIÓN DE LA EXPRESIÓN DE BCL7A TIENE UN EFECTO SUPRESOR TUMORAL	90
5.10 BCL7A PARTICIPA EN VÍAS DE ACTIVACIÓN DE CÉLULAS B Y EN LA REGULACIÓN DE GENES IMPLICADOS EN EL CICLO CELULAR	93
5.11 EL COMPLEJO SWI/SNF ESTÁ FRECUENTEMENTE MUTADO EN EL SUBTIPO GCB-DLBCL	98

6. DISCUSIÓN.....	105
6.1 LAS MUTACIONES DE <i>SPLICING</i> DE BCL7A IMPIDEN SU UNIÓN AL COMPLEJO SWI/SNF.....	111
6.2 LA RESTAURACIÓN DE LA EXPRESIÓN DE BCL7A EJERCE UN PAPEL SUPRESOR TUMORAL EN MODELOS CELULARES DE DLBCL.....	113
6.3 EL COMPLEJO SWI/SNF EN EL CONTEXTO FISIOPATOLÓGICO DEL CENTRO GERMINAL.....	116
7. CONCLUSIONES.....	119
8. BIBLIOGRAFÍA.....	123
9. DERECHOS DE AUTOR Y PERMISOS.....	137
10. APÉNDICES.....	141
<i>APÉNDICE I. Subunidades de los complejos SWI/SNF en mamíferos</i>	<i>145</i>
<i>APÉNDICE II. Variaciones génicas en BCL7A en nuestra cohorte de pacientes y líneas celulares</i>	<i>147</i>
<i>APÉNDICE III. Líneas celulares de DLBCL</i>	<i>149</i>
<i>APÉNDICE IV. Resultados obtenidos por espectrometría de masas.....</i>	<i>151</i>
<i>APÉNDICE V. Artículo de Tesis y artículos adicionales publicados durante la Tesis doctoral.....</i>	<i>155</i>

ÍNDICE DE FIGURAS

FIGURAS INTRODUCCIÓN

Figura 1. La reacción de centro germinal	8
Figura 2. Moduladores transcripcionales que dirigen la reacción de centro germinal	10
Figura 3. Clasificación funcional de los complejos remodeladores de la cromatina dependientes de ATP	13
Figura 4. Ensamblaje por etapas de las subunidades de los distintos complejos SWI/SNF en mamíferos: BAF, PBAF o ncBAF	15
Figura 5. Evolución del complejo SWI/SNF desde levaduras a vertebrados, pasando por <i>Drosophila</i>	16
Figura 6. Mutaciones en determinadas subunidades de los complejos SWI/SN están asociadas con ciertos tipos de cáncer	18
Figura 7. Niveles de expresión de BCL7A en células B maduras	24

FIGURAS RESULTADOS

Figura 8. Análisis de expresión de <i>BCL7A</i> en una colección de 29 líneas celulares hematológicas	55
Figura 9. Curvas de disociación aberrantes obtenidas mediante RT-qPCR en líneas celulares de DLBCL	57
Figura 10. Representación de la secuencia de <i>BCL7A</i> a nivel de ADNc	57
Figura 11. Amplificación de <i>BCL7A</i> mediante PCR para secuenciación de Sanger	58
Figura 12. Amplificación de la región codificante completa de <i>BCL7A</i> a nivel de ADNc	59
Figura 13. Mutaciones con cambio de sentido y mutación puntual en el codón de inicio en líneas celulares de DLBCL	61
Figura 14. Mutaciones de <i>splicing</i> en líneas celulares de DLBCL	62
Figura 15. La línea celular HBL1 presenta una deleción del locus del exón 1 de <i>BCL7A</i>	63
Figura 16. Diferentes mutaciones impiden la expresión de la forma silvestre de <i>BCL7A</i> en algunas líneas celulares de DLBCL	67
Figura 17. Las mutaciones de <i>splicing</i> generan un <i>BCL7A</i> mutante que pierde gran parte del exón 1	69
Figura 18. La línea celular Karpas-1106 presenta una duplicación en tándem que afecta a los exones 2 y 3 en un único alelo	70
Figura 19. Representación de los posibles transcritos obtenidos a partir de <i>BCL7A</i> silvestre o del mutante $\Delta 27$ - <i>BCL7A</i>	72
Figura 20. Expresión de <i>BCL7A</i> a nivel de proteína en distintos contextos celulares	74
Figura 21. Expresión de <i>BCL7A</i> a nivel de proteína en líneas celulares derivadas de linfoma de Burkitt o mieloma múltiple	75
Figura 22. Gráfico de mutaciones en <i>BCL7A</i> en cohortes externas de DLBCL	80
Figura 23. Patrón mutacional de <i>BCL7A</i> en cohortes externas de DLBCL e inactivación bialélica	81

Figura 24. Mutaciones en <i>Bcl7a</i> debido a la acción de AID	85
Figura 25. Mutaciones inducidas por AID en secuencias no específicas	85
Figura 26. Correlación entre las mutaciones inducidas por AID en nuestros modelos de ratón y mutaciones en pacientes con DLBCL	87
Figura 27. La pérdida del dominio BCL7A_Nt impide la unión de BCL7A al complejo SWI/SNF	89
Figura 28. La restauración de la expresión de BCL7A induce un fenotipo supresor tumoral <i>in vitro</i>	91
Figura 29. Xenotrasplante de la línea celular OCI-LY1 transducida con las variantes de BCL7A en ratones NSG	93
Figura 30. Análisis de expresión diferencial mediante RNA-Seq en las líneas celulares OCI-LY1 y VAL después de la sobreexpresión de BCL7A silvestre	94
Figura 31. Los genes diferencialmente expresados de forma común en ambos modelos celulares OCI-LY1 y VAL están relacionados con vías de activación de células B	96
Figura 32. Expresión diferencial de determinados genes al restaurar la expresión de BCL7A	97
Figura 33. <i>Waterfall plot</i> de mutaciones en el complejo SWI/SNF en cohortes externas de DLBCL	99

ABC:	Célula B activada (<i>Activated B-Cell like</i>)
ACN:	Acetonitrilo
AID o AICDA:	Deaminasa de citidina inducida por activación (<i>Activation Induced Cytosine Deaminase</i>)
ADNc:	ADN complementario
AML:	Leucemia mieloide aguda (<i>Acute Myeloid Leukemia</i>)
BAM:	Representación binaria de un mapa de alineamiento (<i>Binary Alignment Map</i>)
CG:	Centro Germinal (<i>Germinal Center</i>)
DLBCL:	Linfoma difuso de células B grandes (<i>Diffuse Large B-Cell Lymphoma</i>)
FA:	Formaldehído
FACS:	Clasificación de células activadas por fluorescencia (<i>Fluorescence-Activated Cell Sorting</i>)
FC:	<i>Fold change</i>
FBS:	Suero bovino fetal (<i>Fetal Bovine Serum</i>)
FDR:	<i>False Discovery Rate</i>
FL:	Linfoma folicular (<i>Follicular Lymphoma</i>)
GADPH:	Gliceraldehído-fosfato deshidrogenasa
GCB:	Célula B del centro germinal (<i>Germinal Center B-cell like</i>)
HRP:	Peroxidasa de rábano picante (<i>Horseradish Peroxidase</i>)
LC-MS/MS:	Cromatografía líquida con espectrometría de masas en tándem (<i>Liquid Chromatography with tandem Mass Spectrometry</i>)
NGS:	Secuenciación de nueva generación (<i>Next-Generation Sequencing</i>)
nt:	Nucleótido/s
MOI:	Multiplicidad de infección (<i>Multiplicity of Infection</i>)
OMS:	Organización Mundial de la Salud
p:	p-valor
PCR:	Reacción en cadena de la polimerasa (<i>Polymerase Chain Reaction</i>)
pb:	Pares de bases

PBS:	Tampón fosfato salino
PMBCL:	Linfoma primario mediastínico de células B (<i>Primary Mediastinal B Cell Lymphoma</i>)
PRC	Complejo represor Polycomb (<i>Polycomb Repressive Complex</i>)
qPCR:	PCR cuantitativa (<i>Quantitative PCR</i>)
rpm:	Revoluciones por minuto
RT-qPCR:	PCR con transcriptasa inversa (<i>Reverse Transcription-qPCR</i>)
SDS:	Dodecilsulfato sódico
SDS-PAGE:	Electroforesis en geles de poliacrilamida en presencia de SDS
SNP:	Polimorfismo de nucleótido único (<i>Single Nucleotide Polymorphism</i>)
UTR:	Región no traducida (<i>Untranslated Region</i>)
SWI/SNF:	<i>SWitch/Sucrose Non-Fermentable</i>
WES:	Secuenciación de exoma completo (<i>Whole Exome Sequencing</i>)
WT:	Silvestre (<i>Wild-type</i>)

Algunos términos en inglés, ampliamente utilizados en Biología Molecular y sin una traducción simple al castellano se presentan en cursiva a lo largo del texto.

1. INTRODUCCIÓN

1.1 CÁNCER

Cáncer es un término que engloba a un conjunto de enfermedades caracterizadas por presentar un crecimiento celular descontrolado y que puede originarse en casi cualquier tejido. La carcinogénesis se manifiesta como un proceso complejo de múltiples etapas donde se produce un cúmulo de variaciones genéticas o epigenéticas que dan lugar a la transformación de una célula sana en una célula tumoral. Estas variaciones permiten a la célula tumoral adquirir características particulares que le posibilitan exceder la tasa proliferativa de una célula sana y dar lugar a un tumor o neoplasia. Además de un crecimiento descontrolado, la célula tumoral puede incluso llegar a invadir otros órganos en un proceso conocido como metástasis, siendo ésta la principal causa de muerte en los pacientes con cáncer¹.

En cuanto a la etiología, muchos tipos de cáncer se asocian con factores de riesgo relacionados con el estilo de vida, el ambiente o la exposición a determinadas sustancias. Algunos ejemplos son el consumo de tabaco o alcohol, llevar una vida sedentaria, tener una dieta no saludable o tener una gran exposición a ambientes y sustancias contaminantes. Por otro lado, mientras que un factor como la edad es considerado un factor de riesgo global, hay otros factores que dependen de la riqueza de un país. De este modo, las infecciones crónicas son un mayor riesgo en países con índice de ingresos bajo-medio, mientras que el impacto del tabaco tiene mayor incidencia en países con altos ingresos².

Tal y como hemos mencionado, todas las células tumorales tienen en común una proliferación celular descontrolada independientemente del tejido al que pertenecen. Pero, ¿qué mecanismos permiten a una célula tumoral proliferar de forma descontrolada? A lo largo de los años se han descrito numerosas señas de identidad del cáncer que incluyen tener un potencial de replicación ilimitado, capacidad de evadir la muerte celular programada o evadir las señales de anti-crecimiento, favorecer la angiogénesis para mejorar la irrigación del tumor, tener autosuficiencia en las señales de crecimiento y/o aumentar la capacidad invasiva para favorecer la metástasis³. Más recientemente, se han descrito nuevas capacidades de las células tumorales como la reprogramación metabólica y la evasión del sistema inmune para evitar ser reconocidas⁴. Además, en esta última revisión, los autores sugieren dos nuevos mecanismos importantes por los cuales las células tumorales adquieren todas las capacidades mencionadas. Estos mecanismos son la inestabilidad genómica, que promueve la variabilidad genética, y la inflamación crónica como detonante de la tumorigénesis.

Los diferentes fenotipos que adquiere una célula para dar lugar a un tumor se adquieren debido a una sucesión de alteraciones genéticas y/o epigenéticas que proporcionan una ventaja de crecimiento y, por lo tanto, dan lugar a una selección clonal⁵. Si la alteración genética supone una ventaja para la célula tumoral es debido a que ésta se produce en determinados genes conductores del cáncer (del inglés *cancer driver genes*). De forma simplificada, las mutaciones conductoras que desencadenan la progresión tumoral ocurren en tres tipos principales de genes: protooncogenes, genes supresores de tumores y genes reparadores de ADN.

- **Protooncogenes:** Estos genes tienen funciones biológicas relacionadas con la activación de mecanismos implicados en la proliferación o crecimiento celular⁶. El tipo de mutaciones que suelen adquirir estos genes son de ganancia de función y pueden producir la activación del gen de forma permanente. Las mutaciones en oncogenes suelen tener un efecto dominante y pueden presentarse en heterocigosis.

- **Genes supresores de tumores:** Estos genes tienen funciones biológicas relacionadas con la inhibición de los procesos de proliferación⁷. Las mutaciones que afectan a estos genes suelen dar lugar a una pérdida de función. Además, normalmente tienen un efecto recesivo, lo cual implica que en determinados casos sea necesario la inactivación de ambos alelos para eliminar por completo la actividad del gen⁸. La inactivación bialélica puede producirse por mutaciones inactivadoras o por pérdida de heterocigosidad, debido a la delección de una parte de un cromosoma o por recombinación homóloga aberrante.

- **Genes reparadores de ADN:** Este conjunto de genes se encargan de reparar errores producidos en el ADN durante la replicación errónea o incompleta del mismo o a través de mutaciones inducidas por radiación o agentes químicos⁹. Por tanto, mutaciones en estos genes están asociadas a una mayor inestabilidad genómica y un aumento en la tasa de mutaciones que podrían afectar a otro tipo de genes como los protooncogenes o genes supresores de tumores.

1.2 NEOPLASIAS HEMATOLÓGICAS

Las neoplasias hematológicas se originan a partir de las células sanguíneas. Aunque el término neoplasia se utiliza normalmente para designar a una masa anormal de tejido que puede ser cancerosa (maligna) o no cancerosa (benigna), cuando se habla de neoplasias hematológica se hace referencia a enfermedades malignas. Teniendo en cuenta la localización, las neoplasias hematológicas se han clasificado de forma más tradicional en leucemia o linfoma. Esta diferenciación se basa en que la leucemia es considerada un tumor líquido en el que hay una proliferación incontrolada de células inmaduras en la sangre. En cambio, el linfoma se produce

de manera más localizada en forma de una masa tumoral sólida en algún órgano linfoide y se origina a partir de células en estadios más maduros, normalmente linfocitos. Pero, actualmente la Organización Mundial de la Salud (OMS) estratifica las neoplasias hematológicas en tres grupos primarios basados en el linaje celular: neoplasias mieloides, neoplasias linfoides y neoplasias de histiocitos o células dendríticas¹⁰. Dentro de cada categoría hay numerosas clasificaciones en función del grado de diferenciación celular, del tipo celular, del nivel de agresividad o de características moleculares referentes a translocaciones o genes particulares.

Las neoplasias hematológicas incluyen un grandísimo número de categorías. A grandes rasgos y de forma simplificada, en la **tabla 1** se muestran las neoplasias hematológicas más frecuentes en la población, así como su incidencia. Como podemos ver, los linfomas son más frecuentes que las leucemias. Dentro de los linfomas, los linfomas de Hodgkin solo representan un pequeño porcentaje frente a los linfomas no Hodgkin, los cuales engloban alrededor de la mitad de todas las neoplasias hematológicas diagnosticadas. El mieloma múltiple, con un porcentaje elevado de incidencia, se clasifica dentro de las neoplasias linfoides derivadas de células B maduras. Más concretamente, el mieloma múltiple se produce por la proliferación anormal de células plasmáticas, que representan el último estadio de diferenciación de las células B.

Tabla 1. Incidencia de las neoplasias hematológicas más frecuentes entre la población

Tipo de Neoplasia Hematológica	Porcentaje por subtipo	Porcentaje
Leucemias	--	30,4%
Leucemia linfoblástica aguda	4%	
Leucemia mieloide aguda	8,7%	
Leucemia linfocítica crónica	10,2%	
Leucemia mieloide crónica	3,7%	
Otras leucemias	3,8%	
Linfomas	--	55,6%
Linfoma de Hodgkin	7%	
Linfoma no Hodgkin	48,6%	
Mieloma múltiple	--	14%

Datos referentes a la población de EE.UU. (años 1975-2006)¹¹

Una particularidad de las neoplasias hematológicas respecto a tumores sólidos es la presencia de translocaciones cromosómicas. Muchas veces, las translocaciones presentes en la célula tumoral son las que condicionan tanto el diagnóstico como el tratamiento a seguir. Así, por ejemplo, dentro de los linfomas no Hodgkin de células B la translocación t(11; 14)(q13; q23) está presente en más del 95% de los linfomas del manto¹². Esta translocación promueve la expresión

del gen de ciclina D1, que sirve como marcador inmunofenotípico en el diagnóstico de este tipo de linfomas. Otro gran ejemplo, es la translocación t(9;22)(q34;q11.2) que sirve de biomarcador diagnóstico en un 95% de los casos de leucemia mieloide crónica¹³. Además, la caracterización de determinadas translocaciones puede derivar incluso en tratamientos concretos. Esto se ha observado en la leucemia promielocítica aguda, un subtipo concreto de leucemia mieloide aguda, que se caracteriza por presentar la translocación t(14;17)(q22;q12). Esta translocación da lugar al gen de fusión *PML-RAR α* que genera un receptor de ácido retinoico truncado. Esta caracterización ha derivado en el uso farmacológico de ácido transretinoico (ATRA), el cual, en conjunto con otros fármacos, ha mejorado el pronóstico de estos pacientes¹⁴. Este es uno de los pocos ejemplos de neoplasias hematológicas que tienen tratamientos moleculares específicos.

1.3 LINFOMA DISFUSO DE CÉLULAS B GRANDES

El linfoma difuso de células B grandes (DLBCL) es el subtipo más prevalente dentro de los linfomas no Hodgkin. Es una forma agresiva de linfoma y representa alrededor de un 35% de todos los linfomas no Hodgkin diagnosticados^{10,15}. Raramente afecta a adultos jóvenes o niños y tiene una edad media de diagnóstico de aproximadamente 70 años. El primer síntoma suele ser el crecimiento rápido de una masa tumoral localizada en zonas ganglionares, acompañada de los denominados síntomas B: fiebre, pérdida de peso y sudoraciones nocturnas. Normalmente se diagnostica *de novo* siguiendo criterios morfológicos o inmunofenotípicos. Pero, en determinados casos el DLBCL aparece como una transformación de algún linfoma más indolente como linfoma folicular (FL) o leucemia linfocítica crónica¹⁶.

A diferencia de otras neoplasias hematológicas, en DLBCL no hay translocaciones relacionadas con su etiología o diagnóstico de forma inequívoca. Más bien, el DLBCL se caracteriza por tener una gran heterogeneidad, donde la adquisición de sucesivas alteraciones génicas termina desencadenando el fenotipo maligno de la enfermedad^{15,17}. Hay determinadas translocaciones que se producen en algunos pacientes de DLBCL y que afectan normalmente a los genes MYC, BCL2 o BCL6. Ahora bien, estas translocaciones no originan genes de fusión como ocurre en leucemias agudas, sino que los genes involucrados quedan regulados por regiones promotoras heterólogas. En consecuencia, se produce una expresión ectópica o exacerbada de genes como MYC, BCL6 y/o BCL2 que pasan a actuar como oncogenes¹⁸.

Finalmente, la frase que mejor define el DLBCL a nivel biológico es la siguiente: “el DLBCL es una expansión maligna de clones de células B que han quedado atrapados en alguna de las diferentes etapas de diferenciación de una célula B en el centro germinal”. Las siguientes secciones

hablarán de diferentes aspectos encaminados a descifrar el significado e implicaciones de esta afirmación.

1.3.1 Biología de las células B: reacción de centro germinal

Primero, para poder comprender con detalle los mecanismos moleculares que desencadenan la linfomagénesis en el DLBCL, es necesario entender la propia biología de las células B, también llamadas linfocitos B. Las células B tienen un papel central en la inmunidad adaptativa humoral a través de la generación de los anticuerpos necesarios para protegernos frente a microorganismos extracelulares y sus toxinas¹⁹. Las células productoras de anticuerpos específicos frente a determinados antígenos son las células plasmáticas o las células B de memoria y éstas se generan a partir de la diferenciación de los linfocitos B maduros (**Figura 1**). Ahora bien, cuando una célula B reconoce un antígeno puede iniciar caminos diferentes. Una opción es que las células B se diferencien a células plasmáticas sin requerir de la acción de los linfocitos T²⁰. La otra opción, dependiente de la acción de los linfocitos T, es que las células B entren en un proceso de maduración en el interior del centro germinal (CG)²¹.

Los CGs son estructuras que se forman transitoriamente en respuesta a un antígeno. Estas entidades se desarrollan en los órganos linfoides periféricos como el bazo, ganglios linfáticos, amígdala o placas de Peyer. El CG se caracteriza por generar un microambiente totalmente especializado donde los linfocitos B completan su ciclo de maduración por afinidad²². Básicamente, el CG sirve para que una célula B activada por antígeno sea capaz de producir anticuerpos con una afinidad mucho mayor que la original por ese mismo antígeno. El nombre de CG procede de la primera vez que fueron caracterizadas histológicamente estas estructuras, ya que se distinguían como regiones con una alta proliferación celular²³. Ya por entonces, se definieron dos zonas dentro del CG, la zona oscura y la zona clara. La zona oscura es donde los linfocitos B *naive*, que acaban de encontrarse con el antígeno por primera vez, se diferencian a centroblastos y proliferan rápidamente. A continuación, migran hacia la zona clara donde pasan a llamarse centrocitos y son seleccionados por afinidad con ayuda de otros tipos celulares como los linfocitos T o las células foliculares dendríticas (**Figura 1**).

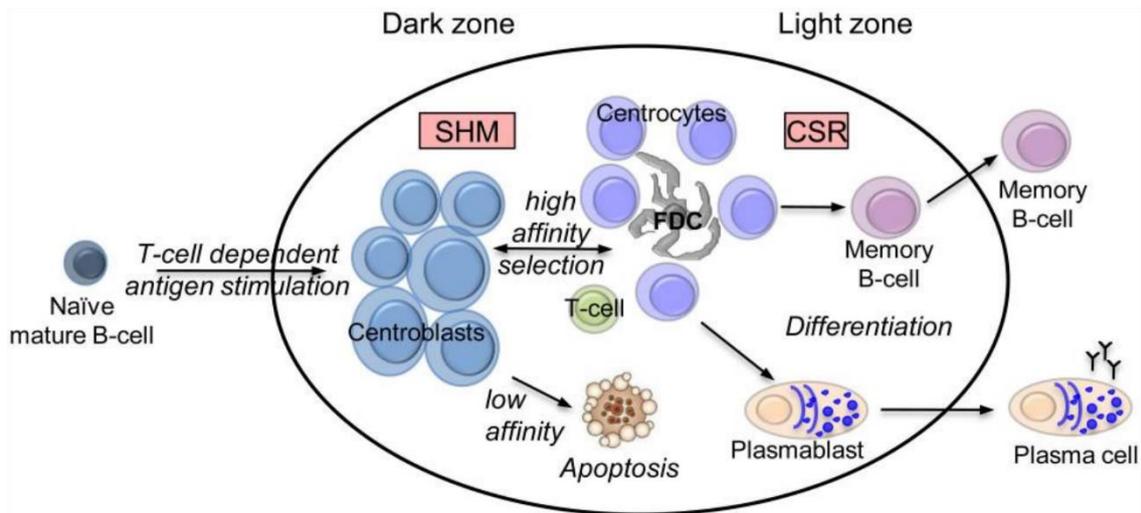


Figura 1. La reacción de centro germinal. Tras la estimulación por antígeno dependiente de células T, las células B *naive* migran a los órganos linfoides secundarios, se diferencian en centroblastos y proliferan en la zona oscura de los CGs. Dentro de la zona oscura, los centroblastos experimentan el proceso de SHM, que introduce variaciones en las regiones variables del locus de las cadenas pesadas y ligeras de las inmunoglobulinas, con el objetivo de aumentar su afinidad por el antígeno. Posteriormente, los centroblastos se mueven a la zona clara, donde se diferencian a centrocitos y pueden experimentar el proceso de CSR. Las células T y las células dendríticas foliculares (FDC) ayudan a los centrocitos a reencontrarse con el antígeno. De esta manera, las células con un receptor de inmunoglobulina con baja afinidad se eliminan por apoptosis, mientras que un subconjunto de centrocitos con alta afinidad por el antígeno se seleccionan para diferenciarse a células B de memoria o a células plasmáticas. El paso de zona oscura a zona clara es bidireccional y un centrocito con afinidad por antígeno puede volver a la zona oscura para seguir proliferando y mejorando su afinidad. Imagen adaptada de Schneider, C. et al. (2011)²⁴.

Hay una serie de procesos que son intrínsecos a la maduración por afinidad que se produce en los CGs. Estos procesos son la hipermutación somática (SHM) y la recombinación de cambio de clase (CSR). La hipermutación somática es el mecanismo por el cual se modifican las regiones variables de los genes de inmunoglobulinas mediante la introducción de mutaciones. Esto permite generar un amplio repertorio de anticuerpos que puedan mejorar la afinidad contra el antígeno²⁵. En cuanto a la recombinación de cambio de clase, es el proceso por el cual se produce un cambio de la región pesada de las inmunoglobulinas para producir anticuerpos de diferentes isotipos que mejoren la respuesta efectora del sistema inmune contra un patógeno determinado²⁶. Los mecanismos de SHM y CSR están orquestados por la enzima Deaminasa de Citosina Inducida por Activación (AID o AICDA)²⁷. La enzima AID tiene la capacidad de deaminar las citosinas para convertirlas en uracilos. Estos cambios, si no se reparan, dan lugar a mutaciones, o bien, pueden generarse roturas de doble cadena que son necesarias para el proceso de CSR.

Los mecanismos de SHM o CSR, aun siendo herramientas imprescindibles para generar respuestas inmunes específicas, representan una espada de doble filo. La generación de roturas

en el ADN puede inducir determinadas translocaciones cromosómicas²⁸. Por otro lado, además de genes de inmunoglobulinas, la enzima AID está directamente relacionada con mutaciones en muchos otros genes. Un claro ejemplo de esta hipermutación somática “aberrante” se ha caracterizado en el locus de *BCL6*, donde mutaciones inducidas por AID en regiones promotoras provocan la desregulación de la expresión de este gen²⁹.

1.3.2 Regulación transcripcional de la reacción de CG

El hecho de que una célula B sea seleccionada para emprender su camino a través del CG es algo excepcional. Una célula B no lleva escrito en su ADN los pasos que debe tomar para realizar el proceso de maduración por afinidad. Por el contrario, determinados cambios a nivel de regulación transcripcional hacen posible que las células B con mayor afinidad por el antígeno migren hacia ciertas zonas de los folículos linfoides y de comienzo la reacción de CG.

Uno de los reguladores transcripcionales más importante para iniciar la reacción de CG es *BCL6*, cuya expresión es rápidamente inducida cuando una célula B es seleccionada por afinidad³⁰. *BCL6* actúa como represor transcripcional y dirige la inhibición de diferentes genes³¹. Algunos de ellos están relacionados con la mayor capacidad proliferativa que adquieren los centroblastos^{32,33}, mientras que la inhibición de otros genes predispone a las células B a tener una mayor permisividad hacia los daños producidos en el ADN durante los procesos de SHM o CSR³⁴. La expresión de genes como *BCL6* se mantendrá activa mientras las células B no terminen su maduración. Eso sí, una vez que las células B son seleccionadas para diferenciarse a una célula productora de anticuerpos, la expresión de *BCL6* se apaga y otros genes como *PRDM1* o *XBP1* toman el control de la regulación transcripcional en la etapa final de diferenciación de las células B. En la **figura 2** se puede ver el patrón de expresión de algunos de los genes implicados en la regulación transcripcional a lo largo de la reacción de CG.

En definitiva, la reacción de CG es un proceso que requiere de una regulación transcripcional muy precisa donde hay muchos genes implicados. Esta complejidad también implica un talón de Aquiles que aprovechan las células tumorales para adquirir ventajas proliferativas cuando alguno de los múltiples genes sufre variaciones genéticas de ganancia o de pérdida de función. De hecho, cerca del 80% de linfomas de tipo no Hodgkin, incluyendo el linfoma de Burkitt, FL o DLBCL, derivan de células B del propio CG (**Figura 2**). Todos los linfomas derivados del CG presentan genes de inmunoglobulinas mutados, indicativo de que han experimentado la reacción de CG³⁶ o incluso mantienen todavía activos los mecanismos de SHM³⁷.

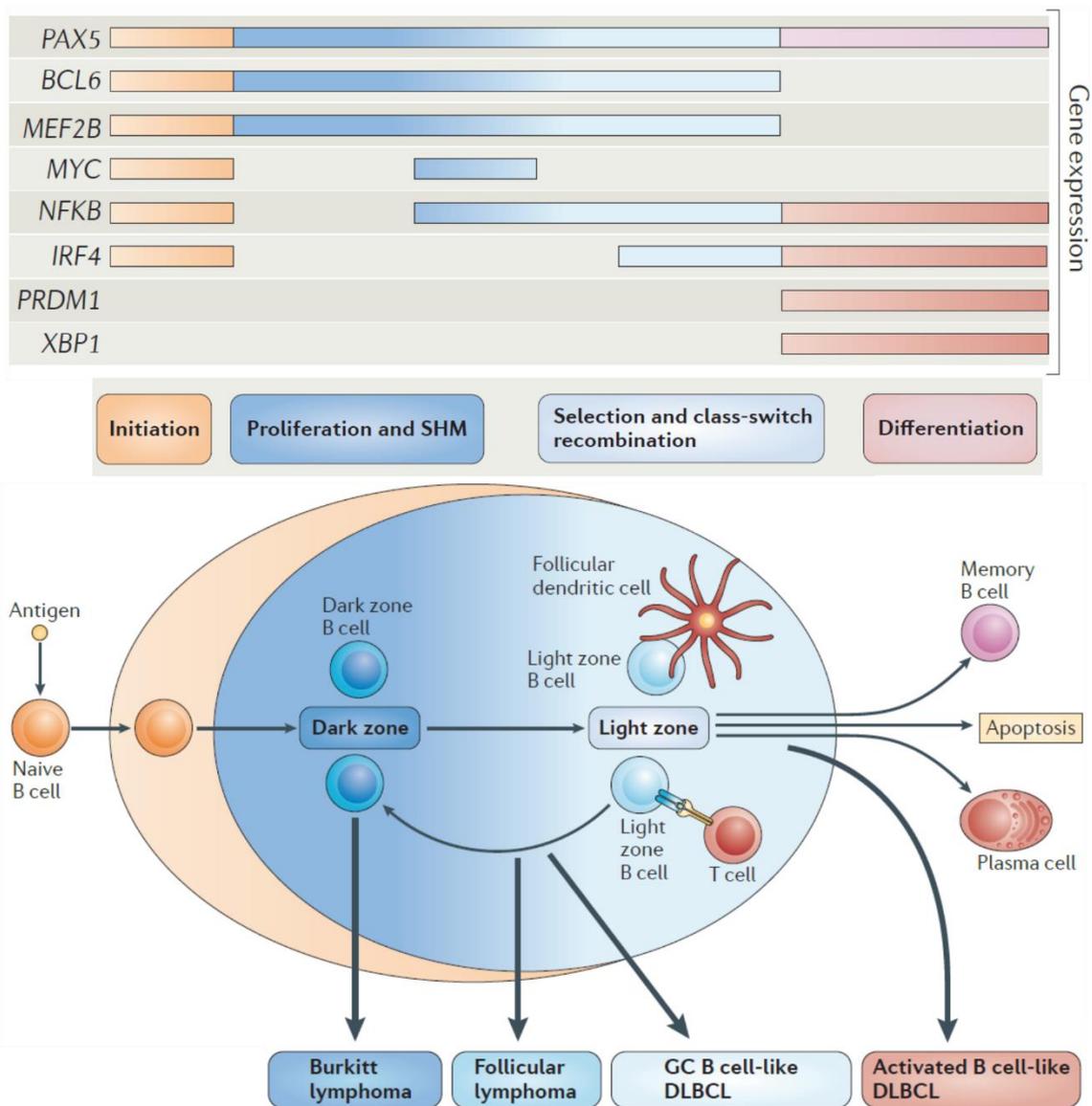


Figura 2. Moduladores transcripcionales que dirigen la reacción de centro germinal. El inicio de la reacción de CG depende de algunos factores de transcripción como BCL6 o MEF2B, cuya expresión está restringida a las etapas que suceden dentro del CG. Otros factores como MYC o IRF4 se expresan durante la iniciación y en posteriores etapas, pero su expresión se apaga momentáneamente en células B que proliferan en la zona oscura. Finalmente, otros genes como *PRDM1* o *XBP1* dirigen la última etapa de diferenciación de una célula B a célula plasmática. Distintos tipos de linfoma que derivan del CG representan clones de células que se asemejan molecularmente a células B atrapadas en distintos estadios de maduración. Imagen adaptada de Basso, K. and Dalla-Favera, R. (2015)³⁵.

1.3.3 Clasificación molecular del DLBCL

El uso de plataformas como los *microarrays* a principio de los años 2000, permitió clasificar los linfomas DLBCL en dos grandes grupos, atendiendo a similitudes con la supuesta célula de origen: (i) célula B de centro germinal, (GCB-DLBCL); (ii) célula B activada, (ABC-DLBCL)^{38,39}. Estos subtipos son indistinguibles morfológicamente pero el estudio de los perfiles de expresión génica ha permitido clasificarlos en función a la similitud que comparten con células B que representan distintas etapas de diferenciación a lo largo del CG. El subtipo GCB es el más

frecuente y es aquel en el que los perfiles de expresión génica se asemejan a los mostrados por los centroblastos o centrocitos. En cambio, el subtipo ABC tiene inhibida la expresión de marcadores típicos del CG en favor de genes como *IRF4* o *XBP1*, que están relacionados con la diferenciación a células plasmáticas⁴⁰. Sin embargo, alrededor del 15% de los pacientes no se pueden clasificar en ninguno de los subtipos mencionados^{38,41}.

Hay un “tercer” subtipo de DLBCL denominado linfoma mediastínico primario de células B grandes (PMBCL). Su incidencia es baja, representando únicamente alrededor de un 6% de todos los DLBCL. En un principio, el PMBCL se consideró como un subtipo de DLBCL con la particularidad de que se originaba a partir de células B localizadas en el timo. Los primeros estudios de expresión génica mostraron que compartía patrones de expresión con el linfoma de Hodgkin⁴² pero a diferencia de este sí que presentaba expresión de genes típicos de células B maduras⁴³. En el año 2001, la OMS lo consideró como una entidad única de linfoma, pero hay autores que incluso a día de hoy siguen mencionando al PMBCL como un subtipo de DLBCL^{39,44}.

La clasificación GCB/ABC se realizó a principios de los años 2000 gracias a plataformas de análisis de expresión génica como los microarrays y su validez se mantiene a día de hoy. Ahora bien, la posterior aplicación de la secuenciación masiva en genomas tumorales de DLBCL ha revelado una sorprendente heterogeneidad a nivel molecular. El número de genes que presentan mutaciones en DLBCL es muy elevado⁴⁵⁻⁴⁷. Esta gran heterogeneidad genera incluso ambigüedades a la hora de definir los genes frecuentemente mutados en DLBCL en función de la cohorte de pacientes analizada⁴⁸. Aun así, dentro de esa heterogeneidad, existen patrones mutacionales representativos en función de la célula de origen. En el subtipo ABC predominan mutaciones en genes implicados en la señalización a través del receptor de la célula B (BCR) y miembros de la señalización por NF- κ B^{49,50}. Por el contrario, genes modificadores de la cromatina se encuentran frecuentemente mutados en el subtipo GCB y en otros linfomas derivados del CG, como es el FL⁵¹. Por tanto, una mejor caracterización molecular de grupos de pacientes es necesaria para entender los procesos biológicos que contribuyen al fracaso de los tratamientos convencionales.

1.3.4 Tratamiento

A día de hoy el tratamiento estándar de primera línea para tratar el DLBCL es la combinación R-CHOP. Las siglas **CHOP** atienden a una combinación de agentes quimioterapéuticos: ciclofosfomida (**C**), clorhidrato de doxorubicina (**H**idroxidaunorrubicina), vincristina (**O**ncovin) y prednisona (**P**). El tratamiento CHOP se estableció como tratamiento de primera línea después de mostrar una menor toxicidad en comparación a otros regímenes quimioterapéuticos⁵².

Posteriormente, en el año 1997, se aprobó el uso de Rituximab (**R**) que es un anticuerpo monoclonal contra el receptor CD20, expresado normalmente en las células B. La inclusión del rituximab en el tratamiento supuso una gran mejora en el pronóstico de los pacientes con DLBCL⁵³⁻⁵⁵.

A pesar de las mejoras en los tratamientos, aproximadamente un 40% de los pacientes recaen o desarrollan enfermedad refractaria y terminan sucumbiendo a la enfermedad⁵⁶. Otros regímenes terapéuticos o combinaciones de quimioterapéuticos son investigados, pero sin mejoras adicionales hasta el momento^{57,58}. Los distintos subtipos de DLBCL en base a los perfiles de expresión génica parecen tener distinta respuesta a los tratamientos actuales, siendo el subtipo ABC el que presenta un mayor riesgo de recaída^{59,60}. En cuanto al tratamiento del subtipo PMBCL, al ser un subtipo que afecta a adultos jóvenes, se han llevado a cabo estudios en los que se ha aumentado la dosis de agentes quimioterapéuticos o se han incluido ciclos de radioterapia⁶¹, habiendo obtenido mejoras de supervivencia en algunos ensayos clínicos⁶².

Hoy en día, mediante la aplicación de la secuenciación masiva, se conoce una mayor cantidad de genes que se encuentran somáticamente mutados en el DLBCL. Conforme se van conociendo en detalle las vías de señalización que contribuyen al desarrollo tumoral, se abren nuevas oportunidades para avanzar hacia una terapia cada vez más personalizada⁶³. Además, nuevas entidades de DLBCL se definen en función a la coocurrencia de determinadas variaciones moleculares, como es el caso de los denominados linfomas de doble hit. Estos linfomas tienen un pronóstico malo y se caracterizan por presentar translocaciones en el gen *MYC* junto a translocaciones en *BCL2* y/o *BCL6*⁶⁴.

En definitiva, es fundamental definir la implicación sobre el fenotipo tumoral de los distintos genes alterados en DLBCL. Solamente de esta forma se podrán lograr terapias complementarias a la quimioterapia actual y que actúen de forma efectiva frente a las vulnerabilidades generadas en los diferentes pacientes. También, debido a que el DLBCL afecta a la población anciana, muchas veces los tratamientos tienen efectos de comorbilidad y los pacientes no se benefician de tratamientos de quimio o radioterapia. En cambio, la inmunoterapia u otras terapias moleculares dirigidas podrían sortear la alta toxicidad de las terapias actuales que reciben los pacientes de DLBCL.

1.4 REMODELACIÓN DE LA CROMATINA Y COMPLEJO SWI/SNF

1.4.1 Complejos remodeladores de la cromatina

El material genético de cada una de nuestras células se encuentra altamente compactado en el interior del núcleo celular en forma de cromosomas. Se denomina cromatina a la mezcla compactada de ADN, ARN y proteínas que forman nuestros cromosomas. Los nucleosomas son la unidad básica de la cromatina y están formados por un octámero de histonas, alrededor del cual se enrollan unos 146 pares de bases de ADN⁶⁵. La estructura de la cromatina varía de forma dinámica y los complejos remodeladores de la cromatina, junto a enzimas modificadoras de histonas, se encargan de reestructurar el complejo histonas-ADN. Esto es necesario para que el ADN sea accesible durante los procesos de transcripción, replicación o reparación. Las enzimas modificadoras de histonas reconocen y marcan covalentemente residuos específicos en las colas de histonas mediante acetilación, metilación, fosforilación, ubiquitinación u otros procesos⁶⁶. Por el contrario, los complejos remodeladores de la cromatina reconocen estas marcas y, a través de la hidrólisis de ATP, movilizan, reestructuran o expulsan los nucleosomas⁶⁷.

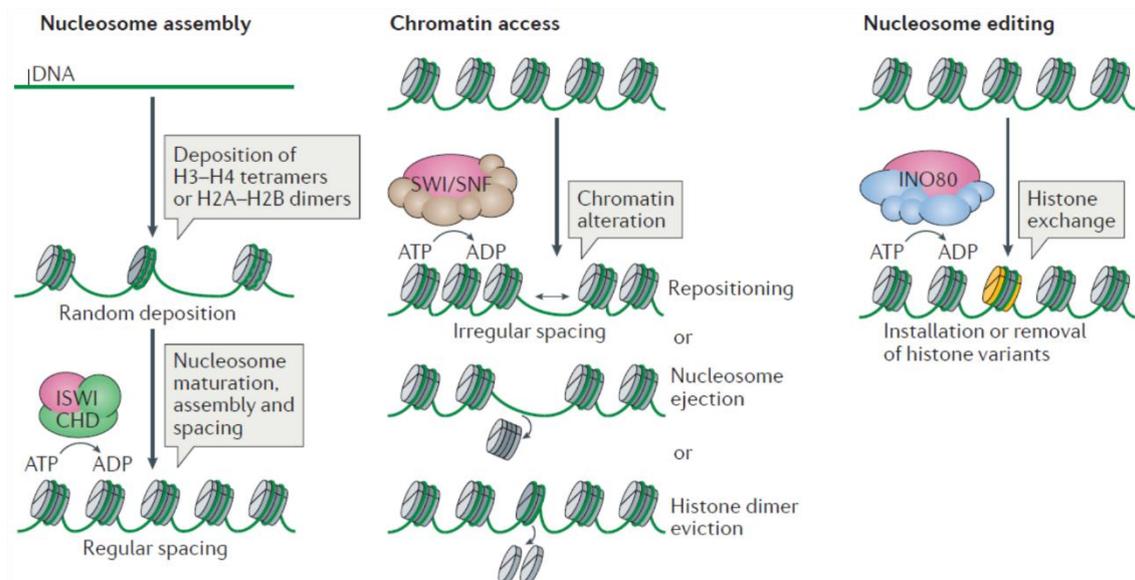


Figura 3. Clasificación funcional de los complejos remodeladores de la cromatina dependientes de ATP.

Los distintos complejos remodeladores están especializados para realizar preferiblemente alguna de las siguientes funciones: organización y ensamblaje de los nucleosomas, controlar la accesibilidad de la cromatina o editar los nucleosomas. Estas funciones son llevadas a cabo preferencialmente por los complejos ISWI-CHD, SWI/SNF e INO80 respectivamente. En cuanto al complejo SWI/SNF, este ejerce una función primaria en controlar la accesibilidad de la cromatina mediante distintos procesos como son repositionar nucleosomas o bien expulsar nucleosomas de forma completa o parcial. Esta clasificación funcional es una simplificación, ya que ciertos subtipos de complejos INO80 o ISWI/CHD son capaces de controlar también la accesibilidad de la cromatina. Imagen adaptada de Clapier, C.R. et al. (2017)⁶⁸.

Los complejos remodeladores de la cromatina dependientes de ATP se clasifican en distintas familias dependiendo de su actividad biológica y de las subunidades que contienen. Estas familias incluyen los complejos SWI/SNF, INO80, ISWI, y CHD/NuRD. Se ha visto que los distintos complejos comparten el mecanismo de translocación del ADN debido a un dominio ATPasa “similar” capaz de realizar esta función⁶⁸. Además, dominios adicionales dentro de la subunidad ATPasa, que varían según la familia, son los que les aportan diferencias en la forma de regular la subunidad ATPasa o unirse a otras subunidades accesorias. En la **figura 3**, podemos ver ejemplos de cómo los diferentes complejos presentan mecanismos específicos a la hora de remodelar el tándem ADN-nucleosoma.

1.4.2 Complejo SWI/SNF: Estructura y función

A lo largo de los años se van describiendo particularidades funcionales de los distintos complejos remodeladores de la cromatina. De entre todos ellos, el complejo SWI/SNF es el que presenta una mayor tasa mutacional en distintos tipos de tumores humanos⁶⁹. Esto ha atraído la atención de la comunidad científica, haciendo que el complejo SWI/SNF sea el complejo remodelador de la cromatina más estudiado hoy en día.

El complejo SWI/SNF es necesario para llevar a cabo distintos procesos biológicos entre los que se incluye la replicación y reparación del ADN^{70,71}, la regulación del desarrollo⁷² o la regulación de la diferenciación celular⁷³. La gran diversidad funcional del complejo SWI/SNF surge de la amplia variabilidad de subunidades que forman estos complejos multiproteicos. A día de hoy, se han descrito complejos SWI/SNF formados por un total de hasta 15 subunidades, que son codificadas por 29 genes⁷⁴. A grandes rasgos, se han descrito tres complejos SWI/SNF diferentes en mamíferos en función de las subunidades que contienen: un complejo canónico BAF (*BRM/BRG-associated factor*); PBAF (*polybromo-associated BAF*); y más recientemente se ha descrito un complejo ncBAF (*non-canonical BAF*)^{75,76} (**Figura 4**). En el **apéndice I** podemos ver una lista de todas las subunidades descritas hasta el momento y los tipos de complejos SWI/SNF que forman.

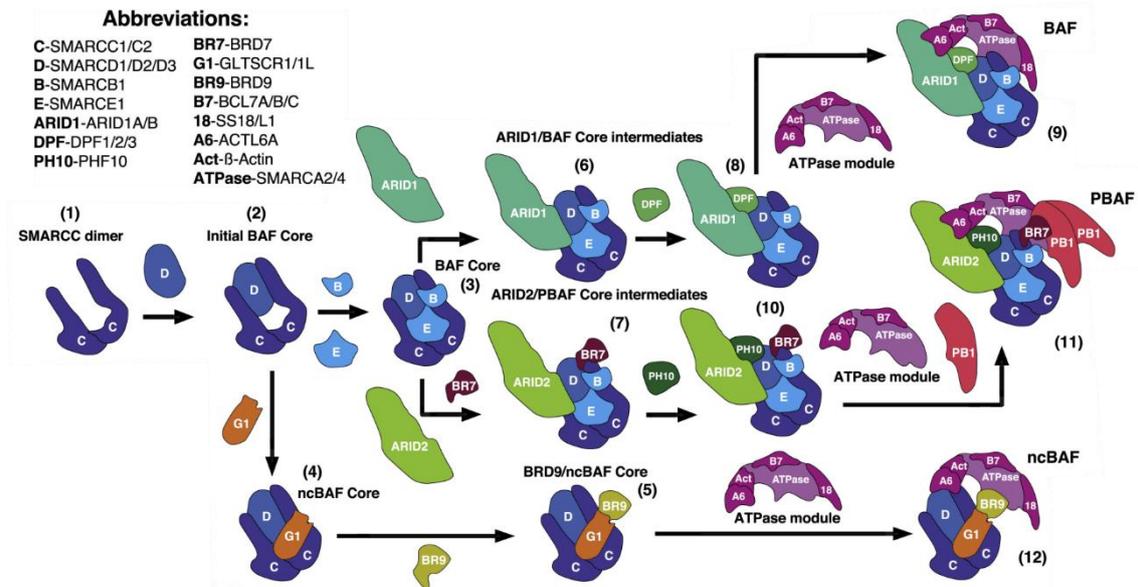


Figura 4. Ensamblaje por etapas de las subunidades de los distintos complejos SWI/SNF en mamíferos: BAF, PBAF o ncBAF. Se muestran las diferentes etapas de ensamblaje de las subunidades del complejo SWI/SNF y las ramificaciones que se producen para construir los diferentes tipos de complejos. Imagen adaptada de Mashtalir, N. et al. (2018)⁷⁷.

La remodelación de la cromatina no sería posible sin la actividad ATPasa de estos complejos. Gracias a diferentes ensayos *in vitro*, se ha visto como la actividad remodeladora de la cromatina recae solamente sobre cuatro subunidades del complejo SWI/SNF⁷⁸. Por un lado, la subunidad ATPasa que puede ser tanto SMARCA4 (también llamado BRG1) como SMARCA2 (BRM) de forma excluyente, y por otro lado subunidades estructurales como SMARCD1 (BAF60A), SMARCC1 (BAF155) y SMARCC2 (BAF170). Estas tres últimas subunidades mencionadas, junto a las subunidades SMARCB1 (BAF47) y SMARCE1 (BAF57), forman lo que se ha denominado núcleo BAF, y está presente en las versiones del complejo BAF o PBAF. En cambio, las subunidades SMARCB1 o SMARCE1 no son necesarias para formar el núcleo ncBAF, el cual se forma a partir de las subunidades GLTSCR1 o GLTSCR1L, que son exclusivas de los complejos ncBAF. Estos núcleos BAF o ncBAF son los primeros en ensamblarse y sirven como soporte para el anclaje del resto de subunidades “accesorias” (**Figura 4**)⁷⁷. Las subunidades “accesorias” aportan a los complejos SWI/SNF distintas funcionalidades a la hora de unirse al ADN, reconocer marcas epigenéticas, servir como anclaje a otras proteínas o contribuir en la regulación transcripcional⁷⁹. Algunas de estas funciones vienen determinadas por los diferentes dominios que presentan algunas de las subunidades, entre los que encontramos bromodominios⁸⁰ (unión a lisinas acetiladas), cromodominios⁸¹ (unión a lisinas metiladas), dominios de unión a ADN y dominios “*PHD finger*” de unión a marcas epigenéticas.

Además, lo que hace dinámico al complejo SWI/SNF en mamíferos es que la composición de subunidades puede incluso variar dependiendo del subtipo celular. A lo largo de la evolución, desde los complejos SWI/SNF en levaduras hasta los complejos en vertebrados, nuevas subunidades han aportado diferentes vías para adaptarse a grandes cambios evolutivos como son la multicelularidad en *Drosophila* o la aparición de la metilación del ADN en los vertebrados (Figura 5)⁸². Un claro ejemplo de cómo las subunidades del complejo SWI/SNF aportan funcionalidades muy específicas dependiendo del subtipo celular, viene dado por la diferente composición de subunidades que se produce durante la diferenciación neuronal en los mamíferos⁸³. En este caso particular, las subunidades PHF10 y ACTL6A, que están presentes en los complejos SWI/SNF de progenitores neuronales (npBAF), son intercambiadas por las subunidades DPF1/2 y ACTL6B en neuronas post-mitóticas. De hecho, cada vez es más evidente que el complejo SWI/SNF tienen importantes funciones en el desarrollo neuronal a intelectual, ya que numerosas discapacidades intelectuales se han relacionado con mutaciones en algunas subunidades del complejo⁸⁴.

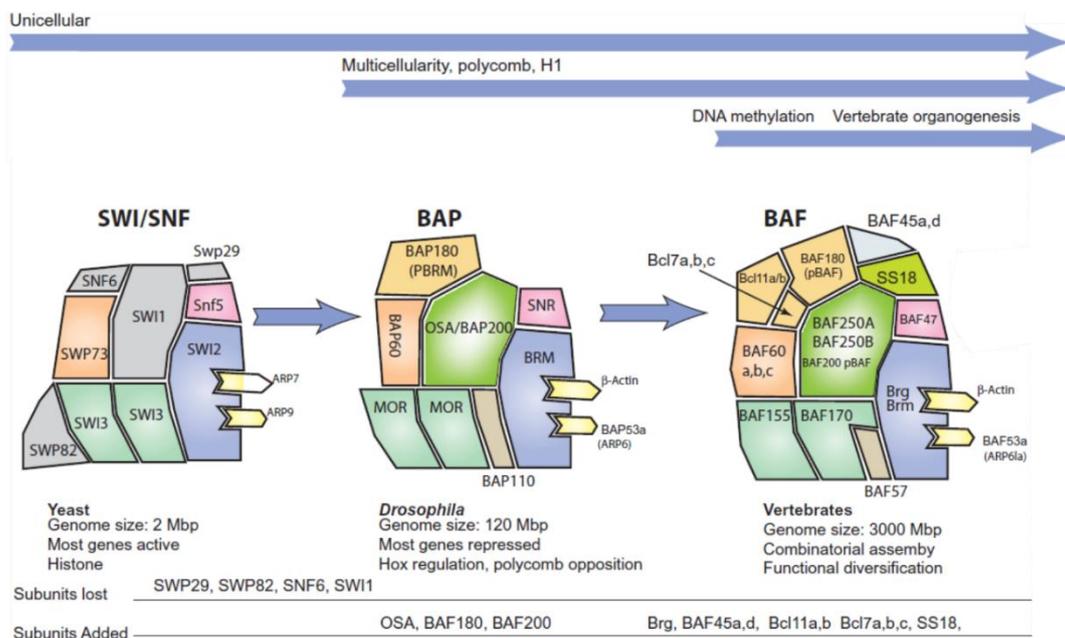


Figura 5. Evolución del complejo SWI/SNF desde levaduras a vertebrados, pasando por *Drosophila*. La composición en subunidades o sus homólogos ha ido variando a lo largo de millones de años de evolución. Los colores usados indican homología entre subunidades. El desarrollo de la multicelularidad y la necesidad de reprimir gran cantidad de genes vino acompañado de la aparición de los complejos represores Polycomb o la histona H1. Estos eventos aparecieron acoplados temporalmente a grandes cambios en la composición de subunidades en los complejos SWI/SNF en levaduras a los complejos BAP en moscas. Posteriormente, eventos como la metilación del ADN, la aparición de genomas más grandes o la complejidad que apareció con los organismos vertebrados vinieron acompañados de una nueva transición en cuanto a subunidades y a la capacidad combinatoria de ensamblaje de los complejos BAF en vertebrados. Imagen adaptada de Kadoch, C. and Crabtree, G. R. (2015)⁷⁹.

Debido a la capacidad del complejo SWI/SNF de movilizar nucleosomas, la función más directa que se le atribuye es la regulación transcripcional. Pero más allá de esa visión simplista, el complejo SWI/SNF también se ha visto asociado a otros procesos celulares. Por un lado, se ha descrito que algunas subunidades como SMARCA4 o ARID1A son necesarias para asistir en diversos procesos relacionados con la reparación del ADN^{85,86}. Además, el complejo SWI/SNF también se ha visto implicado en procesos que ocurren durante la división celular como son la decatenación del ADN o la unión de los telómeros^{87,88}.

En cuanto a la regulación transcripcional, existen evidencias de que el complejo SWI/SNF es capaz tanto de activar como de inhibir la expresión génica⁸⁹. Independientemente de si induce o reprime la expresión de determinados genes, numerosos trabajos han determinado que el complejo SWI/SNF tiene una presencia importante en regiones conocidas como potenciadores (término “*enhancer*” en inglés). Los *enhancers* son regiones de ADN no codificante que sirven como plataforma de anclaje para distintos factores de transcripción que de forma conjunta permiten la regulación de la expresión génica. En este sentido, el complejo SWI/SNF mantendría accesibles estas zonas para permitir la entrada de otros reguladores transcripcionales. Pues bien, se ha visto que el complejo SWI/SNF se asocia preferencialmente a *enhancers* que dirigen la expresión de genes importantes en la diferenciación celular⁹⁰⁻⁹². También, se ha visto cómo el complejo SWI/SNF juega un papel importante en la regulación epigenética de los promotores bivalentes⁹³.

1.4.3 Complejo SWI/SNF y sus implicaciones en cáncer

La primera vez que se relacionó al complejo SWI/SNF con cáncer fue a finales de la década de los 90 cuando se identificaron mutaciones bialélicas en el gen codificante de la subunidad SMARCB1 en prácticamente la totalidad de tumores rabdoideos⁹⁴, un tipo de cáncer pediátrico raro pero muy agresivo. Posteriormente, diferentes estudios *in vivo* corroboraron el papel supresor tumoral de la subunidad SMARCB1 al observar cómo la pérdida de esta subunidad en modelos de ratón originaba tumores similares a los encontrados en humanos^{95,96}. A partir del año 2010, debido al auge de estudios de secuenciación masiva en múltiples subtipos tumorales, se puso de manifiesto la altísima prevalencia de mutaciones de las diferentes subunidades del complejo SWI/SNF en numerosos tipos de tumores^{97,98}.

Ahora bien, no todas las subunidades del complejo SWI/SNF parecen estar mutadas de forma equivalente o ubicua. Mientras que hay subunidades como SMARCE1, BCL7B/C, BAF45A-D o BAF60A-C que apenas se encuentran mutadas en cáncer, hay otras subunidades como ARID1A, ARID1B o SMARCA4 que aparecen mutadas frecuentemente en un amplio espectro de tipos de

tumores⁷⁵. En cambio, hay subunidades del complejo SWI/SNF que presentan mutaciones de forma muy específica en determinados tipos de tumores. Además de las ya mencionadas mutaciones de SMARCB1 en prácticamente la totalidad de tumores rabdoide y en un tipo concreto de meningioma⁹⁹, la subunidad SMARCA4 presenta inactivación bialélica en más del 90% de carcinomas de ovario de células pequeñas de tipo hipercalcémico¹⁰⁰. También, otra subunidad exclusiva de los complejos PBAF, como es PBRM1, presenta mutaciones en más del 40% de los carcinomas renales¹⁰¹.

Además de los tumores rabdoide, otro de los grandes ejemplos en los que variaciones en el complejo SWI/SNF están directamente relacionadas con el desarrollo tumoral son los sarcomas sinoviales. A diferencia de los tumores rabdoide, los sarcomas sinoviales son un caso particular en el cual el complejo SWI/SNF actúa como un oncogén. La particularidad de los sarcomas sinoviales es que presentan la translocación t(X;18), la cual se produce entre la subunidad del complejo SS18 y los genes SSX¹⁰². Esta translocación da lugar a una proteína de fusión oncogénica SS18-SSX, provocando además que la subunidad SMARCB1 deje de ensamblarse en los complejos BAF. Ambos eventos dan lugar a la expresión preferencial de genes relacionados con una mayor proliferación¹⁰³. Finalmente, en la **figura 6** podemos ver la relación entre la prevalencia de mutaciones en subunidades concretas del complejo SWI/SNF con determinados tipos de cáncer.

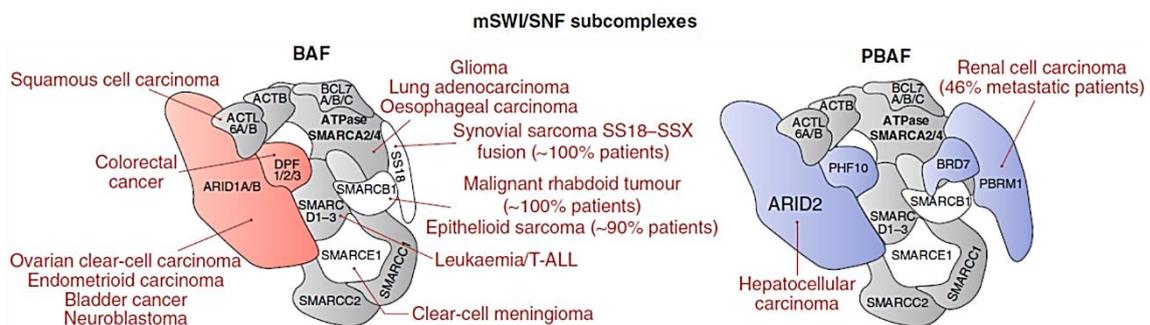


Figura 6. Mutaciones en determinadas subunidades de los complejos SWI/SNF en mamíferos están asociadas con ciertos tipos de cáncer. Las subunidades que son únicas de determinados tipos de complejos SWI/SNF en mamíferos se muestran coloreadas en rojo (BAF) o en azul (PBAF). En esta figura no se muestran los complejos ncBAF, ya que ninguna de las subunidades propias de estos complejos como son GLTSCR1/1L o BRD9 presentan mutaciones de forma específica en algún tipo de cáncer. Imagen adaptada de Valencia, A.M. and Kadoch, C. (2019)¹⁰⁴.

Una vez se ha definido que distintas subunidades del complejo SWI/SNF están mutadas en distintos tipos de tumores, la gran pregunta que surge es ¿cómo contribuyen estas mutaciones al desarrollo tumoral? Tal como hemos comentado previamente, el complejo SWI/SNF está relacionado con la regulación transcripcional y se podría deducir que cualquier alteración que

favorezca la expresión aberrante de un oncogén o la represión de un supresor tumoral podría favorecer la progresión del cáncer. Al mismo tiempo, se ha visto que el complejo SWI/SNF está implicado en procesos de reparación o replicación del ADN^{105,106}. Por consiguiente, mutaciones de pérdida de función de subunidades del complejo SWI/SNF podrían dar lugar a una mayor inestabilidad genómica o a un aumento de la carga mutacional de la célula. Debido a esta dualidad, hay cierto debate en cuanto a si el papel supresor tumoral del complejo SWI/SNF se debe a que los tumores mutantes dan lugar a una regulación transcripcional aberrante, o bien conllevan una inestabilidad genómica asociada. Aunque es difícil de abordar, la teoría de la inestabilidad genómica no parece disponer de evidencias contundentes⁷⁰. Además, se ha demostrado que tumores rabdoideos deficientes en SMARCB1 presentan genomas bastante simples y diploides¹⁰⁷. En cambio, diferentes estudios muestran año tras año cómo defectos en SWI/SNF desencadenan importantes cambios transcripcionales. Por ejemplo, estudios en modelos celulares de sarcomas sinoviales muestran como los complejos BAF dejan de activar la expresión de determinados genes localizados en zonas de *enhancers*. De la misma forma, los complejos BAF favorecen la expresión oncogénica de genes bajo el control de promotores bivalentes¹⁰⁸. Además, la desregulación de programas transcripcionales implicados en la diferenciación celular parece ser una constante en diversos contextos tumorales mutantes para el complejo SWI/SNF⁹⁰⁻⁹².

Finalmente, se puede concluir que modelos tumorales específicos como los rabdoideos o sarcomas sinoviales, donde se sabe que los defectos en el complejo SWI/SNF son el desencadenante de la transformación tumoral, han servido durante estos últimos años para realizar grandes avances en comprender cómo influye el complejo SWI/SNF en la tumorigénesis. En cambio, en tumores más heterogéneos, donde la carga mutacional en las subunidades del complejo SWI/SNF queda diluida y entrecruzada con el amplio espectro de mutaciones en muchos otros genes, es difícil determinar la contribución real de las diferentes subunidades del complejo SWI/SNF en el desarrollo tumoral. Aun así, comprender las consecuencias moleculares de variaciones en determinadas subunidades del complejo SWI/SNF en distintos contextos tumorales tiene un gran valor, ya que podría derivar en el desarrollo de terapias dirigidas que mejoren el pronóstico de determinados pacientes.

1.4.3 Aproximaciones terapéuticas en tumores mutantes para SWI/SNF

Ahora que sabemos que las mutaciones en el complejo SWI/SNF son muy frecuentes en distintos tipos de tumores, la pregunta más inmediata es ¿qué estrategias terapéuticas pueden ser efectivas contra los tumores que presentan estas mutaciones? Las mutaciones en las

subunidades del complejo suelen ser mutaciones de pérdida de función como mutaciones sin sentido, mutaciones con cambio del marco de lectura u otros eventos como grandes deleciones. Sin embargo, las subunidades mutantes en sí mismas no representan posibles dianas terapéuticas para el uso de fármacos. Es por ello, que el interés actual reside en identificar las vulnerabilidades o dependencias que confieren estas mutaciones a las células tumorales y explotar el potencial de dichas “debilidades” para generar nuevos fármacos⁷⁴.

Un concepto terapéutico que puede aplicarse a los tumores mutantes en SWI/SNF es el concepto de letalidad sintética^{109,110}. La letalidad sintética es la coocurrencia de dos eventos que dan lugar a la deficiencia de la expresión o función de dos genes distintos que conduce a la muerte celular. Este fenómeno de letalidad sintética se aplica a cualquier susceptibilidad génica pero el caso de los tumores mutantes para el complejo SWI/SNF es bastante particular, ya que se ha visto que la letalidad sintética se produce entre genes del propio complejo SWI/SNF. El Dr. Schiaffino, antiguo miembro de nuestro laboratorio, escribió un artículo de revisión al respecto¹¹¹. Por ejemplo, se han demostrado numerosos casos en los que la inhibición de SMARCA2 tiene un efecto anti-tumoral en células mutantes para la subunidad SMARCA4¹¹²⁻¹¹⁴. Cabe mencionar que las subunidades SMARCA4 y SMARCA2 son genes parálogos, muy similares en cuanto a secuencia, y además son mutuamente excluyentes. Esto implica que aquellos complejos SWI/SNF que contienen una de las subunidades catalíticas no contienen la otra. Del mismo modo, el fenómeno de letalidad sintética también se ha identificado para otras subunidades parálogas y mutuamente excluyentes como son ARID1A y ARID1B¹¹⁵. La primera conclusión que se puede sacar de estos estudios es que los complejos SWI/SNF residuales, que contienen las subunidades parálogas a las que presentan mutaciones, son capaces de dirigir el fenotipo proliferativo de las células tumorales. Más recientemente, se ha visto que el efecto de letalidad sintética entre subunidades del complejo no se produce solamente entre subunidades parálogas. También se ha observado que tumores rabdoideos o sarcomas sinoviales, que carecen de SMARCB1 por diversos motivos, parecen ser sensibles a la inhibición de BRD9^{116,117}.

Tras haberse caracterizado que el propio complejo SWI/SNF es capaz de dirigir el fenotipo tumoral, algunas de las aproximaciones terapéuticas se basan en intentar bloquear la función del propio complejo SWI/SNF residual que permanece en las células mutantes. Una de las estrategias que se han explorado consiste en usar inhibidores de bromodominios. Estos dominios están presentes en subunidades como SMARCA2, SMARCA4 o BRD9. Desafortunadamente, el uso de inhibidores de bromodominios dirigidos contra SMARCA2 en células mutantes para SMARCA4 no han mostrado efectos anti-proliferativos, ya que la actividad ATPasa parece tener una mayor relevancia que el propio bromodominio para mantener la

arquitectura del complejo SWI/SNF o mantener su unión a la cromatina¹¹⁸. Posteriormente, inhibidores de la función ATPasa han mostrado un efecto anti-proliferativo en xenotrasplantes de células derivadas de pulmón mutantes para SMARCA4¹¹⁹. Otras terapias con gran potencial son aquellas en las que, en lugar de inhibir la actividad de ciertos dominios, lo que buscan es degradar completamente una proteína mediante lo que se denominan PROTACs (del inglés *Proteolysis Targeting Chimeras*)¹²⁰. Esta terapia se ha probado para degradar ciertas subunidades del complejo y se ha conseguido un efecto anti-proliferativo al degradar SMARCA2 en células mutantes para SMARCA4. Desafortunadamente, la aplicación traslacional de las terapias mencionadas tiene ciertas limitaciones, ya que tanto los inhibidores de ATPasa como los PROTACs no son específicos de SMARCA2, sino que también afectan a la actividad SMARCA4 y, por tanto, podrían generar toxicidad en las células no tumorales.

Por otra parte, se ha observado que el propio complejo SWI/SNF, en situaciones en las que no presenta mutaciones o translocaciones, es capaz de contribuir en el fenotipo tumoral. Esto se ha observado en contextos muy particulares de leucemias mieloides agudas (AML), donde se ha visto que los complejos SWI/SNF de las células leucémicas están preferencialmente ensamblados en torno a la subunidad SMARCA4¹²¹. Además, se ha visto en estos modelos que la actividad de SMARCA4 está relacionada con un aumento de la transcripción del oncogén *Myc*¹²². De la misma forma, al igual que en el caso de tumores mutantes para SMARCB1, se ha visto que los modelos de AML son sensibles a la inhibición de BRD9¹²³. En este caso los inhibidores específicos contra el bromodominio de BRD9 sí que han mostrado efecto anti-proliferativo en líneas celulares derivadas de pacientes con AML¹²³. Mientras que los inhibidores del bromodominio de BRD9 no han mostrado eficacia en líneas celulares derivadas de tumores rabdoideos o sarcomas sinoviales, la degradación química de BRD9 sí que ha mostrado eficacia antitumoral^{116,117}.

En resumen, numerosas estrategias terapéuticas van surgiendo para intentar atacar las “debilidades” de las células tumorales mutantes para el complejo SWI/SNF. Aun así, podemos concluir que no existe una terapia ubicua, sino que las aplicaciones son muy dependientes del contexto y subtipo tumoral. Variaciones en ciertas subunidades del complejo SWI/SNF generan consecuencias moleculares muy particulares dependiendo del tipo celular. Por lo tanto, explorar los diferentes contextos puede ser la clave para encontrar beneficios terapéuticos para determinados pacientes.

1.5 BCL7A

El gen *BCL7A* se describió por primera vez en el año 1996 al estudiar una triple translocación que involucraba a los genes *MYC*, *IGH*¹²⁴ y a un gen desconocido en el cromosoma 12 en la línea celular Wien 133, derivada de un paciente con linfoma de Burkitt¹²⁵. Los autores nombraron *BCL7A* al gen que acababan de caracterizar, utilizando el acrónimo 'BCL' (del inglés *B-cell Lymphoma*) en sintonía a otros genes, como *BCL2* o *BCL6*, que suelen estar implicados frecuentemente en translocaciones en neoplasias hematológicas¹²⁶. Sin embargo, *BCL7A* ha sido menos estudiado que otros genes 'BCL' y su función biológica sigue siendo desconocida.

BCL7A forma parte de la familia de genes BCL7 junto a *BCL7B* y *BCL7C*¹²⁷. Estructuralmente, los tres genes parálogos comparten el dominio amino-terminal con una identidad del 90% en los primeros 51 aminoácidos de la secuencia proteica. Estos aminoácidos forman el dominio BCL_N, que está registrado en las bases de datos de proteínas como Pfam: "PF04714 – *BCL*, *N-terminal conserved region*". Pero lo cierto es que no se ha resuelto la estructura tridimensional de ninguno de los miembros de la familia BCL7 y se desconoce la posible función o funciones que pueden tener tanto el dominio BCL_N como el resto de la secuencia peptídica.

Con anterioridad a la época dorada de la secuenciación masiva, diferentes trabajos en cohortes de pacientes relativamente pequeñas, mostraron cómo *BCL7A* perdía su expresión en determinados tipos de tumores. Así, por ejemplo, se ha observado que *BCL7A* se silencia por hipermetilación del promotor en el 48% de los linfomas cutáneos de células T,¹²⁸ y más concretamente en la entidad conocida como micosis fungoide, donde también se ha visto que la expresión de *BCL7A* es menor con respecto a una entidad más indolente de la enfermedad¹²⁹. Por otra parte, también en el subtipo micosis fungoides, se ha observado que la expresión de *BCL7A* se inactiva mediante delección génica en 9/16 tumores¹³⁰. Por último, en otro estudio independiente que analizó perfiles de expresión en diferentes subtipos de linfomas derivados de células T, una menor expresión de *BCL7A* se asoció a un fenotipo más agresivo¹³¹. Todos estos estudios parecen indicar que la pérdida de *BCL7A* está relacionada con un peor pronóstico, pero las implicaciones funcionales de *BCL7A* en neoplasias derivadas de células T es algo que se desconoce a día de hoy.

1.5.1 BCL7A: asociación al complejo SWI/SNF e historial mutacional

Tal y como hemos comentado anteriormente, la aplicación de técnicas de secuenciación masiva en el estudio de genomas tumorales puso de manifiesto que el complejo SWI/SNF estaba frecuentemente mutado en cáncer. Además, en ese mismo momento, en el año 2013 se describieron nuevas subunidades del complejo SWI/SNF entre las que se encontraban *BCL7A*

junto a sus genes parálogos *BCL7B/C*, además de otras subunidades como *BRD9* o *BCL11A/B*⁶⁹. *BCL7A* ya se había asociado previamente al complejo *SWI/SNF* en otros trabajos independientes, pero consideraban a *BCL7A* y a otras subunidades simplemente como “*BAF-associated factors*”^{89,132}.

Los mismos autores que definieron a *BCL7A* como una subunidad del complejo *SWI/SNF* también mostraron que *BCL7A* tenía un patrón mutacional muy concreto. A diferencia de otros genes, que aparecen mutados en diferentes tipos de tumores, las mutaciones de *BCL7A* se concentraban particularmente en neoplasias hematológicas¹⁰³. Por un lado, se encontraron mutaciones en un 20% de los pacientes analizados con linfomas de tipo no Hodgkin¹³³ o en un 22% de pacientes diagnosticados con mieloma múltiple¹³⁴, aunque en este último estudio las mutaciones definidas realmente se encontraban en regiones no codificantes del gen. El DLBCL es el linfoma no Hodgkin más frecuente entre la población y los primeros estudios centrados exclusivamente en este tipo de linfoma salieron a la luz dos años después. Estos primeros estudios de secuenciación masiva manejaban cohortes relativamente pequeñas de entre 40 a 75 pacientes de DLBCL. La realidad es que *BCL7A* no aparecía reflejado como un gen frecuentemente mutado en estos primeros estudios^{45,47,48}. Aun así, poco después se llegó a la conclusión de que se necesitarían al menos 400 pacientes de DLBCL para poder determinar, con el suficiente poder estadístico, aquellos genes mutados con una frecuencia $\geq 5\%$ ¹³⁵. Más recientemente, los últimos trabajos de secuenciación centrados en DLBCL han analizado cohortes suficientemente numerosas para generar un mapa completo de los genes frecuentemente mutados en este tipo de neoplasia hematológica tan heterogénea^{136,137}. En estos estudios, *BCL7A* aparece como uno de los genes que presenta mutaciones somáticas en DLBCL. De la misma forma, *BCL7A* también mostró una prevalencia mutacional cercana al 20% en una cohorte de 105 pacientes de FL¹³⁸. Aun así, a pesar de que *BCL7A* se encuentra mutado recurrentemente en linfomas derivados del CG, a día de hoy no se ha realizado ningún estudio que profundice en el significado funcional o clínico de estas mutaciones.

1.5.2 La expresión de *BCL7A* parece modularse durante la reacción de CG

Tal como hemos visto anteriormente, el gen *BCL7A* presenta mutaciones en linfomas derivados del CG. También, cabe destacar que la expresión de *BCL7A* parece ser un marcador de células B que se encuentran en plena reacción de CG^{38,139}. Asimismo, la expresión de *BCL7A* depende de la correcta función de *BCL6*, ya que se ha observado cómo la expresión de *BCL7A* disminuye en células B que expresan un *BCL6* mutante con efecto dominante negativo¹⁴⁰. Además de los estudios de expresión génica mencionados, se ha observado mediante inmunohistoquímica que

la expresión de *BCL7A* es especialmente elevada en los linfocitos B dentro del CG, pero su expresión se pierde en las células plasmática que abandonan el CG¹⁴¹.

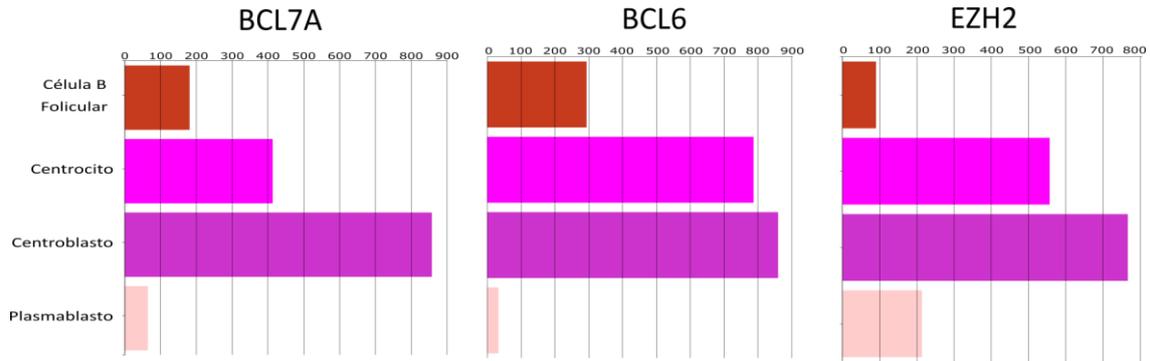


Figura 7. Niveles de expresión de *BCL7A* en células B maduras en comparación con otros genes como *BCL6* o *EZH2* en distintos momentos de diferenciación de una célula B. Estos datos han sido tomados del consorcio *Immunological Genome Project*¹⁴² (<https://www.immgen.org/>).

Hemos hablado en anteriores secciones de cómo la reacción de CG es un proceso en el que hay una regulación transcripcional secuencial por etapas. Hay factores de transcripción como *BCL6* o *PAX5* que son necesarios para dirigir los estadíos iniciales de la reacción de CG. Posteriormente, estos genes se apagan para dar paso a otros genes como *PRDM1* o *XBP1* que permiten la diferenciación de una célula B hacia célula plasmática y su salida del CG³⁵. Siguiendo en esta línea, la expresión de *BCL7A* se modula de forma análoga a genes como *BCL6* o *EZH2* (**Figura 7**). La expresión de todos ellos se induce en centroblastos y centrocitos para posteriormente apagarse en plasmablastos. Ahora bien, mientras que *BCL6* y *EZH2* han sido y son ampliamente estudiados en relación a la biología y la fisiopatología de las células B, las implicaciones de *BCL7A* en estos contextos siguen siendo una incógnita.

Por otro lado, *BCL7A* ha sido caracterizado como uno de los genes que son diana de AID y sufre mutaciones debidas al proceso de SHM^{143,144}. Las mutaciones por AID se producen cerca del sitio de inicio de la transcripción, a menos de 2kb aguas abajo. Además, el proceso de SHM está directamente asociado con altos niveles de transcripción¹⁴⁵. El gen *BCL7A* cumple ambos condicionantes ya que su expresión se induce en células B del CG y su patrón mutacional concuerda con la acción de AID¹⁴⁴. Ahora bien, si las mutaciones generadas por AID favorecen el desarrollo tumoral, es algo que se desconoce.

2. JUSTIFICACIÓN E HIPÓTESIS

En los últimos años, diferentes estudios de secuenciación en genomas tumorales han puesto de manifiesto la alta prevalencia de mutaciones de las distintas subunidades del complejo remodelador de la cromatina SWI/SNF. Además, se han vinculado esas mutaciones al desarrollo tumoral, ya que se ha visto que tanto ratones genéticamente modificados como ratones con mutaciones de pérdida de función en línea germinal para varias subunidades del complejo SWI/SNF tienen predisposición a desarrollar tumores. La novedad de estas apreciaciones deja en desventaja al complejo SWI/SNF en comparación con otros genes supresores de tumores u oncogenes que llevan siendo estudiados durante décadas. Es por ello que el papel del complejo SWI/SNF como supresor tumoral en diversos tipos de neoplasias ha atraído el interés de la comunidad científica. De tal forma, este estudio pretende colaborar en comprender las posibles implicaciones terapéuticas asociadas a mutaciones en genes que codifican para las subunidades del complejo SWI/SNF y así mejorar el pronóstico y tratamiento de estos pacientes.

El complejo SWI/SNF lo forman multitud de subunidades. Algunas de ellas presentan homología con las de los complejos SWI/SNF en levaduras, pero otras, que no están presentes en levaduras, han sido caracterizadas recientemente. Es el caso de *BCL7A*, cuya asociación al complejo se puso de manifiesto en el año 2013. Además, diferentes estudios han mostrado cómo las mutaciones en *BCL7A* parecen restringirse a determinadas neoplasias hematológicas derivadas de células B como los linfomas no Hodgkin.

En la actualidad, la función biológica de *BCL7A* es en gran parte desconocida, aunque diferentes estudios han demostrado que podría desempeñar un papel importante en el CG, donde su expresión se modula durante la diferenciación de una célula B. Por otro lado, *BCL7A* se ha mostrado como uno de los genes que puede sufrir mutaciones a través de un proceso conocido como hipermutación somática aberrante. Este mecanismo se produce en células B durante la reacción de CG. Ahora bien, se desconoce si estas mutaciones son seleccionadas por las células tumorales y confieren una ventaja proliferativa que desencadene el desarrollo tumoral.

Dada la reciente asociación de *BCL7A* al complejo SWI/SNF y el patrón de mutaciones tan específico que presenta, este trabajo se ha centrado en dilucidar las posibles implicaciones de este gen en el desarrollo tumoral de neoplasias hematológicas, donde *BCL7A* podría desempeñar un papel supresor tumoral que se desconoce hasta el momento.

3. OBJETIVOS

El **objetivo general** de este estudio es contribuir a determinar el papel que podría desempeñar el gen *BCL7A* en el desarrollo tumoral. Más concretamente, nos centraremos en abordar las implicaciones de *BCL7A* en neoplasias hematológicas, donde existen la mayoría de indicios que apuntan a que *BCL7A* podría tener un papel relevante.

Objetivos específicos:

1. **Determinar la inactivación de la expresión génica de *BCL7A* en líneas celulares derivadas de neoplasias hematológicas.** Este primer estudio tiene como objetivo definir un modelo de enfermedad en la cual *BCL7A* pudiese tener un papel importante en el diagnóstico, pronóstico y/o tratamiento de los pacientes.
2. **Establecer un modelo funcional para estudiar la actividad de *BCL7A*.** El modelo óptimo de línea celular será aquel en el que la expresión de *BCL7A* se haya perdido debido a alteraciones genéticas o epigenéticas. Para ello estudiaremos las consecuencias a nivel molecular de las variaciones genéticas detectadas.
3. **Restaurar la expresión de *BCL7A* en aquellos modelos celulares que sean defectivos en *BCL7A*.** Estos modelos funcionales nos permitirán:
 - a. Estudiar los cambios fenotípicos tras la restauración de la expresión de *BCL7A* a nivel celular y a nivel de organismo, mediante la realización de xenotrasplantes.
 - b. Estudiar los cambios en los perfiles de expresión génica tras la restauración de la expresión de *BCL7A*.

4. MATERIALES Y MÉTODOS

4.1 MUESTRAS BIOLÓGICAS: LÍNEAS CELULARES Y PACIENTES

4.1.1 Colección de líneas celulares para el estudio piloto

En un primer estudio se utilizó una colección de líneas celulares que englobaban distintos tipos de neoplasias hematológicas (**Tabla M1**). Todas las líneas fueron cultivadas con el objetivo de recoger pellets celulares. Además, se guardaron viales de cada una de las líneas en nitrógeno líquido para su posible uso futuro.

Tabla M1. Líneas celulares utilizadas en el estudio piloto

Línea celular	ATCC	DSMZ	Origen/descripción	Tipo neoplasia
Daudi		ACC 78		Linfoma Burkitt
DB	CRL-2289			DLBCL (GCB)
HBL1			Nozawa, Y. et al., 1988	DLBCL (ABC)
HUT-78			Bunn P A, Jr. et al., 1996	CTCL
JJN-3		ACC 541		Mieloma múltiple
Jurkat		ACC 282		T-ALL
K-562		ACC 10		CML
Karpas-422			Dyer, M. J. et al., 1990	DLBCL (GCB)
Karpas-1106			Nacheva, E. et al., 1994	DLBCL (PMBCL)
KM-H2		ACC 8		Linfoma Hodgkin
KMS-28BM			Inoue, J. et al., 2004	Mieloma múltiple
KU-812		ACC 378		CML
L-363	CRL-9068			Mieloma múltiple
Loucy		ACC 394		CML
MOLT-4	CRL-1582			T-ALL
Namalwa	CRL-1432			Linfoma Burkitt
NB4		ACC 207		APL
OCI-LY1		ACC 722		DLBCL (GCB)
OCI-LY3		ACC 761		DLBCL (ABC)
Raji		ACC 319		Linfoma Burkitt
Ramos		ACC 603		Linfoma Burkitt
REH		ACC 22		Leucemia precursor B
Riva		ACC 585		DLBCL (ABC)
RL		ACC 613		DLBCL (GCB)
RPMI-8226		ACC 402		Mieloma múltiple
RS4:11	CRL-1873			T-ALL
U-266		ACC 9		Mieloma múltiple
U-2932		ACC 633		DLBCL (ABC)
Wien 133			Nacheva, E. et al., 1987	Linfoma Burkitt

APL: Leucemia promielocítica aguda; CML: Leucemia mieloide crónica; CTCL: Linfoma cutáneo de células T; T-ALL: Leucemia linfocítica aguda de célula T.

4.1.2 Panel de líneas celulares de DLBCL

Se utilizó una colección de líneas celulares derivadas de pacientes diagnosticados con DLBCL para secuenciar el gen *BCL7A* a nivel de ADN genómico y realizar así un exhaustivo análisis mutacional sobre la secuencia codificante del gen de interés.

Este estudio agrupó un total de 41 líneas celulares derivadas de pacientes que engloban los distintos subtipos de DLBCL. De estas, 25 derivan del subtipo GCB (DB, DOHH2, Farage, HT, Karpas-422, NU-DHL-1, OCI-LY1, OCI-LY19, OCI-LY4, OCI-LY7, OCI-LY8, Pfeiffer, PR1, RL, ROS-50, SC-1, SU-DHL-4, SU-DHL-5, SU-DHL-6, SU-DHL-10, SU-DHL-16, Toledo, VAL, WSU-DLCL2 y WSU-NHL), 11 del subtipo ABC (Granta-452, HBL1, HLY-1, MD901, NU-DUL-1, OCI-LY3, OCI-LY10, RC-K8, Riva, SU-DHL-8 y U-2932), 2 del subtipo mediastínico (PMBCL) (Karpas-1106 y MEDB1) y tres no clasificadas (Karpas-231, OZ y 380). El material genético relativo a todas estas líneas celulares y los pacientes fue cedido por nuestro colaborador José Ángel Martínez Climent (CIMA, Navarra, España) y no tuvimos que extraer ADN y/o ARN nosotros mismos. En el **apéndice II** podemos encontrar una lista con los nombres y origen de todas las líneas celulares derivadas de DLBCL.

4.1.3 Muestras de pacientes de DLBCL y FL

Para el análisis mutacional de *BCL7A* contamos con ADN extraído de biopsias de 38 pacientes diagnosticados con DLBCL. Algunos de los pacientes habían sido previamente clasificados en función del subtipo de DLBCL¹⁴⁶. En total, había 15 pacientes GCB-DLBCL, 2 pacientes ABC-DLBCL y 21 pacientes DLBCL no clasificados. Además, se incluyeron 15 muestras de pacientes diagnosticados con FL. Para la recogida de muestras, se obtuvo el consentimiento informado de los pacientes en conformidad con la declaración de Helsinki. El estudio fue aprobado por los comités de ética del Consejo Superior de Investigaciones Científicas, la Universidad de Salamanca y la Universidad de Navarra.

4.2 EXTRACCIÓN DE ADN, ARN Y TRANSCRIPCIÓN INVERSA

El ARN total fue extraído a partir de pellets celulares mediante el uso del reactivo Trizol (*TRI Reagent*; Cat#T9424, Sigma-Aldrich; St. Louis, MO, EE. UU.), siguiendo las instrucciones del fabricante. La concentración de ARN y su pureza fueron medidas mediante el uso de un espectrofotómetro *Nanodrop 2000c* (ThermoFisher Scientific, Waltham, MA, EE. UU.). Para la conversión de ARN a ADN complementario (ADNc), se utilizó 1µg de ARN de partida y la transcripción inversa se realizó mediante el kit *RevertAid™* (ThermoFisher Scientific), siguiendo las instrucciones del fabricante.

El ADN genómico utilizado en el análisis de variación de número de copias se extrajo a partir de pellets celulares mediante un protocolo casero que se detalla a continuación. Sobre el pellet celular se añadió 1mL de buffer de extracción (4mM Na₂EDTA, 100mM NaCl, 20mM Tris-HCl; pH 7,5) y se homogeneizó la muestra. A continuación, se añadieron 100µL de una solución acuosa de SDS al 10% y se agitó en vortex durante unos 30 segundos. Luego, se añadieron 10µL de una solución de 20mg/mL de proteinasa K recién preparada y se incubaron las muestras durante toda la noche (máximo 24h) a una temperatura de 55°C. Posteriormente, se dejaron atemperar las muestras y se añadieron 0,8mL de una solución 5,3M NaCl y se agitaron durante 15 segundos. A continuación, se centrifugó durante 5 minutos a 4°C y 13000 rpm. El precipitado blanco se descartó y nos quedamos con el sobrenadante, trasvasándolo a un tubo nuevo. Entonces, se añadió 1 volumen de isopropanol frío, almacenado a -20°C, y se invirtió el tubo unas 5 veces. Para aislar el precipitado de DNA que se forma, centrifugamos a máxima velocidad durante 30 minutos. Posteriormente, se hicieron dos lavados sucesivos con 1mL de etanol al 70% y centrifugación de 30 minutos a máxima velocidad. Por último, se deja secando el ADN durante unos 20 minutos a 37°C y se resuspende el pellet final en agua.

4.3 ANÁLISIS DE EXPRESIÓN DE *BCL7A* EN EL ESTUDIO PILOTO

Se llevó a cabo la optimización de un ensayo RT-qPCR para el análisis de expresión de *BCL7A* utilizando *SYBR Green* (Sigma-Aldrich, Cat-KK4618) y los siguientes cebadores directo (Fw) y reverso (Rv): *BCL7A*-Fw1: 5'-CTCCCCAGCCTCCGTCTCC-3' y *BCL7A*-Rv1: 5'-GTCACCAACGGTCACCCATTT-3'. Las reacciones de qPCR contenían un total de 30ng de ADNc, 5µL de *SYBR Green 2X* y 0,2µL de cada uno de los cebadores (Concentración inicial de 10µM). Las condiciones de PCR fueron de 2 minutos a 50°C, 10 minutos a 95°C, y posteriormente 15 segundo a 95°C y 1 minuto a 60°C durante 40 ciclos. Todas las muestras se midieron en triplicado. Para cada muestra experimental, se determinó la cantidad de *BCL7A* y se utilizó la medida del gen endógeno *GAPDH* para normalizar. La expresión de *BCL7A* y *GAPDH* se calculó mediante el método de curva estándar y se calculó la media del log(cantidad *BCL7A*/*GAPDH*) para cada muestra.

4.4 ANÁLISIS DE VARIACIÓN DE NÚMERO DE COPIAS

Para el análisis de variación de número de copias se utilizaron sondas *TaqMan™* (nº catálogo: 4400291 ID: Hs02932986_cn; ThermoFisher Scientific). El análisis se realizó para las siguientes líneas celulares: Daudi, DB, HBL1, Karpas-1106, KM-H2, OCI-LY1, OCI-LY7, Riva, ROS-50, SU-DHL-4 y VAL. Se realizaron dos repeticiones y cuatro réplicas por repetición. Las reacciones de PCR se realizaron en un sistema de PCR en tiempo real (7000 Real-Time PCR system; Applied

Biosystems) y los resultados se analizaron utilizando el software CopyCaller (v2.1). Los valores estimados de $\Delta\Delta C_t$ para cada línea celular se promediaron entre las dos repeticiones del ensayo. Por último, los valores del número de copia se estimaron tal y como se describe en la guía del usuario de CopyCaller.

4.5 ESTUDIOS DE SECUENCIACIÓN EN LÍNEAS CELULARES Y COHORTE PROPIA DE PACIENTES DERIVADOS DE DLBCL Y FL

El gen *BCL7A* fue secuenciado a nivel de ADN genómico en 41 líneas celulares y 38 tumores primarios de pacientes con DLBCL. Además, se incluyeron en el estudio 15 muestras de pacientes diagnosticados con FL. Distintas parejas de cebadores fueron diseñadas para amplificar la secuencia codificante de *BCL7A*, incluyendo los seis exones y sus respectivas secuencias de *splicing* adyacentes (**Tabla M2**). Las distintas reacciones de amplificación se llevaron a cabo en un volumen final de 20 μ l usando 1 U de polimerasa Taq (Cat#D4545, Sigma-Aldrich) y las siguientes condiciones de PCR: una primera etapa a 94°C durante 1 minutos, seguida de 40 ciclos con etapa de desnaturalización a 94°C durante 10 segundos, etapa de anillamiento a 65°C durante 30 segundos y etapa de extensión a 72°C durante 30 segundos, y una etapa final a 72°C durante 10 minutos. En el caso concreto de la amplificación del exón 1, se tuvo que realizar un programa específico de PCR para mejorar el rendimiento y la especificidad (**Figura M1**). Esto se debe probablemente a la presencia de secuencias repetidas en la región 5'-UTR.

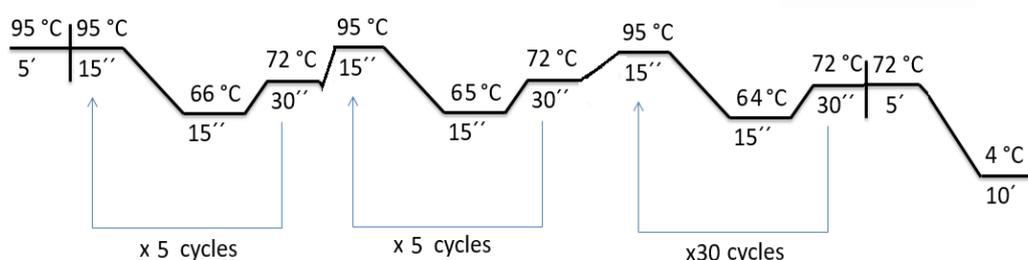


Figura M1. Esquema del programa de PCR utilizado para amplificar el exón 1 de *BCL7A*.

Los productos de PCR se separaron mediante electroforesis en geles de agarosa al 1,5 (%m/v) y, tras cortar la banda, éstos fueron purificados mediante un kit (*GenElute™ Gel Extraction Kit*, Cat#NA1111, Sigma-Aldrich) siguiendo las recomendaciones del fabricante. Posteriormente se realizó la secuenciación mediante el método de Sanger, a través de la compañía StabVida (Setúbal, Portugal), que hacen uso del kit Big Dye versión 3.1 y analizan las muestras en equipos 3730 XL de la compañía Applied biosystems (Foster city, CA, EE. UU.).

Tabla M2. Lista de cebadores utilizados para la amplificación de los diferentes exones de *BCL7A* a nivel de ADN genómico.

Exón	Cebador directo (5'-3')	Cebador reverso (5'-3')	Tamaño amplicón (pb)
1	GCTCTGTGCAGCTGCCGCC	GCTCTCTGGAGCCGCGCTCG	397
2	CCAAGGGAGTGTGGCCTGTG	CAGCACGGTGGGTAGCAGTG	248
3	CTCTGAGCACGTGTGCGCAC	GACCGAGTGCTTGCCAAGG	225
4	GCATTGAGGCACAGGGATGC	GCACATGACAGGGCCTCCTG	355
5	GCCTCAAAGACCACTGGCC	CAAGGCAGAGCCCAGGACTG	416
6	CTTGGCTGACCTTCGGCCTC	CACCCAAGAGGAATCGCTGG	248

4.6 ESTUDIOS DE SECUENCIACIÓN EN COHORTES EXTERNAS DE PACIENTES DE DLBCL

4.6.1 Descarga de datos

Para realizar los análisis mutacionales en cohortes externas, primero se descargaron los archivos BAM de secuenciación de exomas completos (WES), obtenidos de los siguientes estudios:

(i) Reddy et al. (2017), descargados del *European Genome-Phenome Archive* (EGAD00001003600). Filtramos los archivos BAM según una región genómica alrededor de *BCL7A* (chr12:122,460,090-122,460,091, versión del genoma humano hg19). Se excluyeron 13 archivos de WES debido a que tenían una cobertura baja o nula en la región de *BCL7A* mencionada, pese a que las muestras excluidas no tenían deleciones reportadas que afectaran al locus de *BCL7A*. Esto dejó un tamaño de muestra efectivo de 988 muestras de DLBCL.

(ii) Schmitz et al. (2018), estudio al que se accedió a través de la base de *Genotypes and Phenotypes* (dbGaP) del *National Institutes of Health* (NIH), conjuntos de datos phs0014444.v1.p1, phs000178.v10.p8 y phs001175.v2.p2. En total, se recopilamos 553 archivos BAM de WES utilizando la siguiente consulta en el Portal de datos *Genomic Data Commons* (GDC) (<https://portal.gdc.cancer.gov>):

```
cases.project.project_id in ["CTSP-DLBCL1", "NCICCR-DLBCL", "TCGA-DLBC"]
and cases.samples.sample_type in ["Primary Tumor", "Tumor"] and
files.data_category in ["Sequencing Reads"] and
files.experimental_strategy in ["WXS"]
```

Los 553 archivos BAM se descargaron a través de la interfaz de programación de aplicaciones (API) de GDC utilizando la opción "BAM slicing" ("?gencode-BCL7A"). Los archivos de mutaciones halladas por los autores del estudio original se obtuvieron del sitio web de la publicación en GDC (<https://gdc.cancer.gov/about-data/publications/DLBCL-2018>).

4.6.2 Detección y anotación de variantes

Debido a la falta de datos de muestras normales pareadas para la mayoría de los pacientes, se realizó un análisis mutacional no pareado de los archivos BAM descargados, restringiéndolo al gen *BCL7A*. Para la detección de variantes, se utilizó BCFtools (v1.7; www.htslib.org) utilizando los mismos genomas de referencia que se habían utilizado para generar los archivos BAM. Después de la detección de variantes, se aplicaron filtros estrictos para eliminar falsos positivos (QUAL > 60, FMT/DP ≥ 8):

```
bcftools mpileup -B -f ${reference_fasta} -Q 13 -q 1 -a "FORMAT/DP" -O  
u ${bam} | bcftools call -mvO u | bcftools filter -e "%QUAL<60 ||  
FMT/DP<8" > out.vcf
```

```
bcftools merge -m none *vcf | bcftools norm -m -any -f ${reference_fasta}  
> merged.vcf
```

A continuación, las coordenadas genómicas de todas las variantes se convirtieron a la versión hg38 del genoma humano utilizando el paquete de R 'liftOver' (versión R 3.5.2, Bioconductor versión 3.8). Por último, se anotaron las variantes utilizando ANNOVAR (versión 2017-07-17;¹⁴⁷) y las siguientes bases de datos: ensGene (v20170912), 1000g_2015aug_all, exac03, avsnp150 y dbnsfp33a. Se eliminaron todas las variantes que no afectaban a exones o sitios de *splicing*, así como aquellas que tenían una frecuencia poblacional superior al 1% en el proyecto *1000 Genomes* o en *ExAc*. Después de comprobar la concordancia entre las mutaciones exónicas detectadas por nosotros y por los autores originales, se añadieron las mutaciones del sitio de *splicing* a la lista de mutaciones exónicas reportada por los autores originales. Finalmente, se realizó una inspección manual de todas las muestras mutadas en *BCL7A* mediante el software *Integrative Genomics Viewer* (IGV), con el fin de confirmar las nuevas mutaciones detectadas, hallar posibles mutaciones no detectadas por los métodos utilizados y, en el caso de pacientes con dos o más mutaciones, evaluar si estas se encontraban en el mismo alelo o en alelos distintos.

4.7 ESTUDIO DE LA IMPLICACIÓN DE AID EN EL PATRÓN MUTACIONAL DE BCL7A

Este estudio fue realizado en colaboración con el laboratorio *B cell Biology*, dirigido por la Dra. Almudena R. Ramiro. Se trata de un laboratorio experto en el estudio de la biología de las células B y, en particular, en el estudio de los mecanismos propios de la fisiología del CG como es la hipermutación somática. En el año 2018 publicaron un artículo definiendo una estrategia para analizar a gran escala genes que sufren mutaciones debido al proceso de hipermutación somática¹⁴⁸. La librería generada en su estudio no incluyó el análisis del gen *BCL7A*. Por tanto,

decidimos analizar el perfil mutacional de *Bcl7a* en las muestras utilizadas previamente en su estudio.

Estas muestras se generaron mediante el aislamiento de ADN genómico procedente de células B del CG (Cd19⁺Fas⁺GL7⁺), a partir de las placas de Peyer de ratones *Ung*^{-/-} *Msh2*^{-/-} y *Aicda*^{-/-}, tal y como se describe en el artículo original¹⁴⁸. Posteriormente, se amplificó mediante PCR una región que cubría la 5'-UTR y el primer exón de *BCL7A*. Se utilizaron 100ng totales de ADN y se usaron los siguientes cebadores: Directo (Fw) 5'-GCTTCCCAGAGTGAGGTCC-3'; Reverso (Rv) 5'-GCTGCACAATGGAGCTC-3'. Las reacciones de amplificación se llevaron a cabo en un volumen final de 25µl usando 1 U de la polimerasa *Phusion High-Fidelity* (Cat#F530S, ThermoFisher Scientific) y las siguientes condiciones de PCR: una primera etapa a 98°C durante 2 minutos, seguida de 26 ciclos con etapa de desnaturalización a 98°C durante 5 segundos, etapa de anillamiento a 57°C durante 30 segundos y etapa de extensión a 72°C durante 30 segundos, y una etapa final a 72°C durante 10 minutos. Tras varios intentos, vimos que la eficiencia de la PCR era muy baja, lo cual se debía probablemente al gran contenido de citosinas y guaninas en la región amplificada. Este inconveniente se solucionó gracias al uso de formamida al 5% (%v/v) como suplemento en la reacción de amplificación. Los productos de PCR se purificaron tras la electroforesis en gel de agarosa al 1% (*Illustra Gel Band Purification kit*; GE Healthcare, Chicago, IL, EE. UU.) y posteriormente se cuantificaron haciendo uso del equipo Qubit Fluorometric (ThermoFisher Scientific).

Posteriormente, los productos de PCR se fragmentaron usando un sonicador y las librerías se prepararon siguiendo las instrucciones del fabricante (*NEBNext Ultra DNA Library Prep*; New England Biolabs, Ipswich, MA, EE. UU.). La secuenciación se llevó a cabo en la plataforma HiSeq 2500 (Illumina, San Diego, CA, EE. UU.). Las lecturas se realizaron de un único extremo (*single-end*) de fragmentos de una longitud de 61pb y las lecturas se alinearon mediante el software *Novaling* al genoma de referencia GRCm38.75. Los alineamientos se procesaron haciendo uso de samtools 1.9. Posteriormente, el análisis mutacional se realizó utilizando un script Perl personalizado, generado por Ángel Álvarez Prado¹⁴⁸. Solo los nucleótidos cubiertos por más de 100 lecturas que pasaron el umbral de calidad Q20 (es decir, la precisión de la lectura >99%) tanto en el modelo *Ung*^{-/-} *Msh2*^{-/-} como en *Aicda*^{-/-} se consideraron para el análisis. Eso supuso un total de 312pb que se corresponden con las coordenadas genómicas chr5:123344457-123344769. Las diferencias en los recuentos totales de mutaciones entre ratones *Aicda*^{-/-} y *Ung*^{-/-} *Msh2*^{-/-} se evaluaron mediante un test de Fisher de dos colas seguido de corrección mediante FDR.

4.8 CULTIVOS CELULARES

Las líneas celulares tumorales se cultivaron a 37°C en una atmósfera al 5% de CO₂. La línea celular OCI-LY1 se cultivó en medio IMDM (Cat-201980-032, ThermoFisher Scientific) mientras que las líneas celulares HBL1, K-562 y VAL se cultivaron en medio RPMI (Cat-L0498-500, Biowest, Riverside, MO, EE. UU.). Cada cultivo celular se complementó con un 10% de suero bovino fetal (FBS) (Cat-10270-106, Gibco™ ThermoFisher Scientific) y 100 U/ml de estreptomycin y penicilina (Cat-P0781-100ML, Sigma-Aldrich). Para cada una de las líneas celulares se realizó un análisis para comprobar la posible contaminación por micoplasma. Para ello, se utilizó el kit Venor®GeM-qEP (Minerva Biolabs, Berlín, Alemania).

4.9 GENERACIÓN DE VECTORES LENTIVIRALES

Los plásmidos empaquetadores usados en este estudio están basados en el virus de la inmunodeficiencia humana (VIH). El plásmido empaquetador psPAX2 y el plásmido de la envuelta pMD2.G, ampliamente usados en la bibliografía¹⁴⁹, han sido amablemente proporcionados por el laboratorio de Didier Trono (Números de referencia de Addgene #12260 y #12259 respectivamente). Por un lado, el plásmido empaquetador psPAX2 codifica para los genes *gag*, *pol*, *tat* y *rev*. Por otro lado, el plásmido pMD2.G codifica para la proteína G del virus de la estomatitis vesicular (VSV-G). En cuanto al vector lentiviral pLVX-IRES-ZsGreen1 (Cat#632187, Clontech, Shiga, Japón), éste contiene el promotor del virus humano citomegalovirus (más concretamente contiene el “promotor temprano del citomegalovirus humano (P_{CMV IE})”) y expresa el fluorocromo ZsGreen1 a través de un transcrito bicistrónico junto con el posible transgén clonado. El vector pLVX-IRES-ZsGreen1 también contiene el elemento WPRE (*Woodchuck posttranscriptional regulatory element*) y el fragmento central de polipurina (cPPT). Finalmente, el plásmido lentiviral pUltra-Chili-Luc fue cedido por el laboratorio de Malcom Moore (Addgene #48688) y ha sido usado para la expresión estable de luciferasa en los ensayos de bioluminiscencia *in vivo*.

El esqueleto del plásmido pLVX-IRES-ZsGreen1 se usó para clonar tanto la forma silvestre de BCL7A como las variantes Δ 27-BCL7A (isoformas corta y larga). Los diferentes insertos fueron generados mediante PCR, haciendo uso de cebadores que contenían las dianas para las enzimas de restricción XbaI y BamHI. Estas dianas estaban adyacentes a los codones de iniciación (ATG) y terminación (TGA) respectivamente. Posteriormente, los insertos y el plásmido receptor fueron digeridos doblemente con las enzimas XbaI/BamHI y la ligación se llevó a cabo con la enzima ligasa T4 (Cat#EL0011, ThermoFisher Scientific) mediante incubación a 16°C durante toda la noche.

El esqueleto del plásmido SE-SAR2(LTR) se usó del mismo modo para clonar tanto la forma silvestre de BCL7A como las variantes Δ 27-BCL7A (isoformas corta y larga). El uso alternativo de este vector lentiviral se realizó debido a la baja eficiencia, bajo nivel de expresión y alto silenciamiento de las partículas lentivirales generadas a partir del vector pLVX-IRES-ZsGreen1 en la línea celular VAL. El plásmido SE-SAR2(LTR) expresa el transgén a través del promotor SFFV (*Spleen Focus Forming Virus*) y contiene el elemento SAR2¹⁵⁰ insertado dentro de la secuencia 3`LTR, con el objetivo de mejorar los niveles de expresión y reducir el silenciamiento génico. A continuación, para generar las diferentes construcciones SE-SAR2(LTR) expresando las diferentes variantes de BCL7A junto con el fluorocromo ZsGreen1, se digirieron los plásmidos pLVX-BCL7As-IRES-ZsGreen1 con las enzimas SpeI/PstI. Después, los diferentes fragmentos conteniendo las variantes BCL7A/IRES/ZsGreen1/WRPE fueron clonados en los plásmidos SE-SAR2(LTR) previamente digeridos con las enzimas BamHI/SbfI. De esta forma, se obtuvieron los plásmidos finales denominados SAR2(LTR)-BCL7As-IRES-ZsGreen1. De forma similar, el plásmido lentiviral control SAR2(LTR)-IRES-ZsGreen1 se construyó usando como donador el plásmido pLVX-IRES-ZsGreen1.

4.10 PRODUCCIÓN Y TITULACIÓN DE LAS PARTÍCULAS LENTIVIRALES

La producción de partículas lentivirales se llevó a cabo usando células HEK293T como células empaquetadoras y la transfección de estas células se llevó a cabo mediante lipofección. Las células HEK293T se crecieron en medio DMEM (Cat#L0103-500, Biowest) suplementado con FBS al 10% y penicilina-estreptomicina al 1%. El día previo a la transfección, las células se plaquearon en placas de Petri de 10 cm de diámetro (Cat#SIAL0167, Sigma-Aldrich) a una confluencia en torno al 95%. Previamente a la transfección, se renueva el medio dejando un volumen final de 5mL. En ese momento se procede a preparar la mezcla de transfección. En un volumen de 0,5mL de medio sin FBS se hace la mezcla de plásmidos (18 μ g totales de ADN) conteniendo el plásmido lentiviral junto con el plásmido empaquetador psPAX2 y el plásmido de la envuelta pMD2.G en proporciones 3:2:1 respectivamente. Una vez tenemos el conjunto de plásmidos, estos se mezclan a su vez con 45 μ L de LipoD293 (Cat#SL100668, Signagen Laboratories, Rockville, MD, EE. UU.) previamente diluidos en 0,5mL de medio sin FBS y se incuba la mezcla durante 20 minutos a temperatura ambiente. Se añade la mezcla de plásmidos/LipoD293 sobre las células y tras 5 horas de incubación se retira la mezcla para posteriormente añadir unos 7mL de medio completo cuidadosamente sobre las células. Los sobrenadantes, conteniendo las partículas lentivirales, se recogieron 48 y 72 horas después y se pasaron a través de filtros de 0.45 μ m de diámetro (Cat#FPE404030, JET Biofil, Guangzhou, China) para, posteriormente, preparar alícuotas y guardarlas rápidamente a -80°C.

Para todas las transducciones realizadas con alguno de los plásmidos que expresan el fluorocromo ZsGreen1, el porcentaje de células transducidas fue determinado en base al incremento de fluorescencia de dicho fluorocromo, mediante análisis por citometría de flujo. El cálculo del título viral de las partículas lentivirales (ifu/mL: unidades infectivas por mililitro) se realizó en base al porcentaje de células ZsGreen1⁺ detectadas en un rango lineal de una dilución seriada del sobrenadante. Para el cálculo del título viral se utilizó la línea celular K-562 que tiene una gran permisividad para ser transducida.

4.11 ENSAYOS DE CRECIMIENTO CELULAR COMPETITIVO

Los ensayos de crecimiento celular competitivo han sido llevados a cabo para estudiar el efecto en la proliferación de la restauración de la expresión de BCL7A. Esta técnica, a diferencia de otras, permite detectar cambios fenotípicos sutiles¹⁵¹. El uso de plásmidos lentivirales que expresan de forma estable el fluorocromo ZsGreen1 permite la selección, mediante FACS, de aquellas células que han sido transducidas. De esta forma, mediante el uso de un cultivo mixto que contiene tanto células transducidas ZsGreen1⁺ como células no transducidas ZsGreen1⁻, y a través del uso de citometría de flujo, se pueden hacer medidas de la relación de células ZsGreen1⁺/ ZsGreen1⁻ y relacionarlo con la tasa de proliferación celular.

Las líneas celulares OCI-LY1 y VAL se transdujeron con las construcciones SAR2(LTR)-BCL7As-IRES-ZsGreen1 expresando BCL7A de tipo silvestre, el mutante Δ 27-BCL7A o el vector vacío. Las células fueron transducidas tanto con las isoformas cortas como largas de las variantes BCL7A silvestre y mutante Δ 27-BCL7A. Las células se transdujeron mediante 3 ciclos de infección consecutivos con un MOI=2 para ambas líneas celulares. Posteriormente, las células ZsGreen1⁺ transducidas se aislaron mediante FACS (BD FACSAria, BD Biosciences, San Jose, CA, EE. UU.). A continuación, se mezclaron con células no transducidas (ZsGreen1⁻), también aisladas por FACS. A partir de la mezcla, se mantuvieron cultivos mixtos que contenían ambos tipos de células y los cultivos fueron analizados por citometría de flujo para medir la abundancia de la población ZsGreen1⁺ a lo largo del tiempo (BD FACSVerser, BD Biosciences). El crecimiento celular se dedujo del porcentaje de células ZsGreen1⁺ que expresan BCL7A silvestre, Δ 27-BCL7A o el vector vacío y se normalizó teniendo en cuenta el porcentaje de células ZsGreen1⁺ a tiempo inicial de un día después de la mezcla para cada una de las condiciones.

4.12 XENOTRASPLANTE DE LA LÍNEA CELULAR OCI-LY1 EN RATONES NSG

El estudio de bioluminiscencia in vivo, mediante el uso de xenotrasplantes, se realizó en estricta conformidad con las recomendaciones de la guía para el cuidado y el uso de animales de

laboratorio elaborada por el comité bioético de la Universidad de Granada. De la misma manera, los protocolos fueron aprobados por la Comisión de ética de experimentación animal de la Universidad de Granada. Todos los procedimientos se realizaron bajo anestesia por inhalación de isoflurano y se hizo todo lo posible para minimizar el sufrimiento de los animales.

Se utilizaron 30 ratones NOD.scid Il2rg tm1Wjl/SzJ (NSG) de ocho semanas de edad, facilitados por el Servicio de Experimentación Animal de la Universidad de Granada. Se utilizaron solamente ratones macho para evitar posibles efectos hormonales. Los xenotrasplantes se realizaron con células derivadas de la línea celular OCI-LY1. Primero, se generó una línea celular OCI-LY1 que expresaba de forma estable la enzima luciferasa. Para ello, la línea celular OCI-LY1 se transdujo con partículas lentivirales (MOI=1) que contenían la construcción pUltra-Chili-Luc. Posteriormente, las células transducidas se purificaron por FACS gracias a la expresión del fluorocromo dTomato. Una vez que teníamos la línea celular OCI-LY1/Luc⁺, esta se transdujo con 3 tipos de lentivirus diferentes que contenían las construcciones pLVX-IRES-ZsGreen1 para expresar BCL7A silvestre, el mutante $\Delta 27$ -BCL7A, o bien el vector vacío (10 ratones/grupo). En este caso, las células transducidas se purificaron por FACS gracias a la expresión del fluorocromo ZsGreen1 y se dejaron crecer en cultivo hasta obtener el número necesario.

Cada ratón se inyectó, por vía intravenosa a través de la cola, con 5×10^6 células OCI-LY1/Luc⁺ que expresaban las 3 construcciones mencionadas. Posteriormente, se inyectó D-luciferina a diferentes tiempos a los que se tomaron las imágenes de bioluminiscencia. La D-luciferina se inyectó por vía intraperitoneal disuelta en PBS a una dosis de 150mg/kg de peso corporal. Después de 5 a 8 minutos, los animales fueron anestesiados en la cámara oscura usando 3% de isoflurano en el aire a 1,5 L/min y O₂ a 0,2 L/min/ratón, y las imágenes fueron tomadas en una cabina conectada a una cámara (IVIS, Xenogen, Alameda, CA, EE. UU.). El tiempo de exposición fue de 3 minutos en la opción *large binning*, y la cuantificación de la emisión de luz se realizó en fotones por segundo utilizando el software Living Image (Xenogen). El sistema IVIS adquiere una imagen fotográfica bajo luz blanca y mide al mismo tiempo la señal bioluminiscente de forma cuantitativa. Posteriormente, el sistema es capaz de superponer la señal bioluminiscente, expresada en fotones por segundo, sobre la imagen creando una imagen final que se muestra como un mapa de intensidad. Las imágenes se tomaron en los días 0, 7, 14 y 21 después de inyectar las células y se tomaron con los ratones tanto en posición ventral como dorsal.

4.13 ANÁLISIS DE EXPRESIÓN GÉNICA MEDIANTE RNA-SEQ

El ARN se extrajo utilizando el kit *RNeasy Mini* (Qiagen, Hilden, Alemania), siguiendo las instrucciones del fabricante. El ARN se aisló de células OCI-LY1 y VAL, positivas para el reportero

ZsGreen1, que habían sido previamente separadas por FACS tras ser transducidas con las construcciones SAR2(LTR)-BCL7As-IRES-ZsGreen1, que expresaban BCL7A silvestre, Δ 27-BCL7A o el vector vacío. Las células fueron transducidas mediante 3 ciclos de infección consecutivos con un MOI=2. Se analizaron dos réplicas biológicas de transducciones independientes. Los pellets celulares se tomaron 10 y 17 días después de la transducción para las líneas celulares OCI-LY1 y VAL respectivamente. La concentración y calidad del ARN extraído se midieron utilizando el fluorómetro Qubit 4 (ThermoFisher Scientific) y el Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, EE. UU.). Las bibliotecas de ARNm se prepararon utilizando 1 μ g de ARN de partida y el kit de preparación de bibliotecas *TruSeq Stranded mRNA* (Illumina). Las bibliotecas de ARNm se secuenciaron en el sistema NextSeq 500 (Illumina) utilizando el modo *highest output* y las longitudes de lectura de 75pb de extremo emparejado (*paired-end*).

Se realizó un control de calidad de los archivos FASTQ mediante el uso del software FastQC (v0.11.8). Las lecturas se alinearon con el genoma humano de referencia GRCh38.d1.vd1 (descargado de <https://gdc.cancer.gov/about-data/data-harmonization-and-generation/gdc-reference-files>) usando STAR (versión 2.6.1b) en el modo *two-pass*. El software GENCODE v31 se usó como referencia para la anotación y, posteriormente, los archivos BAM *multi-sample* se separaron en archivos BAM *single-sample*. La calidad de los archivos BAM se llevó a cabo haciendo uso de Picard (v2.20.4) y QualiMap (v2.2.1) en combinación con multiqc (v1.7). También se confirmó que la contaminación del ARN ribosómico era menor al 1% y se usó el software IGV para confirmar la presencia, a nivel de ARN, de las mutaciones esperadas según el gen BCL7A exógeno introducido en cada una de las muestras. El número de lecturas asignadas a cada una de las réplicas biológicas fue de aproximadamente 70 millones.

Los números de lectura por gen (*read counts per gene*) se obtuvieron utilizando htseq-count (v0.11.2) con las siguientes opciones: "-m intersection-nonempty -f bam -t exon -i gene_id -r pos -s reverse". Se realizaron análisis posteriores utilizando R (versión 3.6.1 y versión de Bioconductor 3.9) y los paquetes 'edgeR', 'DESeq2' y 'EnhancedVolcano'. Se realizaron clústeres jerárquicos y análisis de componentes principales para confirmar las similitudes entre las réplicas y las condiciones. DESeq2 se utilizó, con los parámetros predeterminados, para el análisis de expresión diferencial, considerando cada combinación de línea celular y/o tratamiento como una condición diferente. El análisis de enriquecimiento de términos por ontología de genes se llevó a cabo utilizando la herramienta de anotación génica y análisis *Metascape* (<http://metascape.org/gp/index.html>).

4.14 ANÁLISIS DE EXPRESIÓN DE *BCL7A* MEDIANTE RT-QPCR

Se realizó un análisis de RT-qPCR para medir la expresión de *BCL7A* en nuestros modelos de líneas celulares transducidas con las construcciones que expresan las distintas variantes de *BCL7A* o un vector vacío. La detección se realizó utilizando SYBR Green (Sigma-Aldrich, Cat-KK4618) y los siguientes cebadores específicos: *BCL7A*-Fw2: 5'- GTGACACCCTACGAATCTAC-3' y *BCL7A*-Rv2: 5'-CACTTCTCTGTCTCTCTCTCTCT-3'. Un microgramo de ARN se utilizó para la transcripción inversa con el kit *RevertAid RT* (Cat-K1691, ThermoFisher Scientific) utilizando cebadores aleatorios y siguiendo las instrucciones del fabricante. Las reacciones de qPCR contenían un total de 30 ng de ADNc, 5µL *SYBR Green 2X* y 0,2µL de cada cebador (Concentración inicial de 10µM). Las condiciones de qPCR fueron de 2 minutos a 50°C, 10 minutos a 95°C, seguido de 40 ciclos con etapas de 15 segundos a 95°C y 1 minuto a 60°C. Todas las muestras se midieron en triplicado. Para cada muestra experimental, se determinó la cantidad de *BCL7A* y de una referencia endógena (*GAPDH*). La expresión de *BCL7A* y *GAPDH* se calculó mediante el método de curva estándar y se calculó la media del log (cantidad *BCL7A/GAPDH*) para cada muestra.

4.15 WESTERN BLOT

Para el uso rutinario de análisis de proteínas por Western blot se siguió la siguiente metodología. La proteína total se extrajo usando el buffer de lisis RIPA (150mM NaCl, 1% NP-40, 0,5% desoxicolato de sodio, 0,1% SDS y 50mM Tris-HCl; pH 7,5) que contenía inhibidores de proteasas y fosfatasa (0,2mM PMSF, 7mM OV₄ and 1X *cOmplete Mini EDTA-free Protease Inhibitor Cocktail*). La concentración de proteína total (mg/µL) fue calculada usando el reactivo Bradford (Cat#M172, VWR, Radnor, PA, EE. UU.) y usando la proteína BSA para la recta patrón. Los Western blot se desarrollaron en condiciones desnaturizantes mediante el uso de buffer de carga que contenían agentes reductores y detergentes. Las proteínas se resolvieron mediante electroforesis en geles SDS-PAGE al 10% de poliacrilamida. Posteriormente, las proteínas fueron transferidas a membranas de PVDF y se bloquearon las membranas usando una disolución de leche libre de grasa al 5% (% m/v) en PBS con 0.1% (v/v) de Tween. La proteína *BCL7A* se visualizó mediante el uso de un anticuerpo policlonal (#HPA019762, Sigma-Aldrich). En cuanto a la proteína p21 se usó el siguiente anticuerpo (12D1: Cat# 2947, Cell Signaling, Danvers, MA, EE. UU.). La proteína Tubulina-α (sc-23948, Santa Cruz Biotechnology, Santa Cruz, CA, EE. UU.) se usó como control de carga. Las diluciones de todos los anticuerpos primarios mencionados fueron 1:1000 en una disolución de leche libre de grasa al 3% (% m/v) en PBS con 0.1% (v/v) de Tween. Posteriormente, tras lavar las membranas con PBS 0.1% Tween, estas se incubaron con

los anticuerpos secundarios específicos de especie (anti-conejo HRP, dilución 1:2000 o anti-ratón HRP, dilución 1:1000; Dako/Agilent Technologies). Las imágenes de Western blot se adquirieron mediante técnicas de quimioluminiscencia y un sistema de cámara ImageQuant LAS 4000 (GE Healthcare). Para la cuantificación de bandas se utilizó la herramienta *Plot Lanes* del software ImageJ v1.49c.

Las líneas celulares HAP1 Parental y HAP1 BCL7A *Knockout* (Número de clon 2434-10) (Product ID: HZGHC002434c010, Horizon Discovery, Waterbeach, Reino Unido), se usaron para determinar la especificidad del anticuerpo. En cuanto a la línea HAP1 BCL7A *Knockout* modificada genéticamente, esta contenía una delección de 7pb en el gen *BCL7A* (211_217delAAATGGG) que genera un cambio en el marco de lectura.

En cuanto a los Western blot que reflejan los resultados de ensayos de inmunoprecipitación, la metodología se detalla en el siguiente apartado.

4.16 INMUNOPRECIPITACIÓN Y ESPECTROMETRÍA DE MASAS (LC-MS/MS)

La línea celular HBL1 fue transducida con las construcciones pLVX-IRES-ZsGreen1 expresando la forma silvestre de BCL7A o el mutante $\Delta 27$ -BCL7A (solamente la isoforma corta en ambos casos). Las células fueron transducidas en tres ciclos consecutivos usando un MOI=1 para las distintas condiciones. Pasados cinco días, las células fueron separadas por FACS para obtener una población pura ZsGreen1⁺ y se cultivaron hasta obtener un número de células adecuado. Posteriormente, se extrajo la proteína total y se cuantificó siguiendo la metodología indicada en la sección Western blot. Para la inmunoprecipitación, se partió de 1mg de proteína total y se incubó durante toda la noche a 4°C con 1µg de anticuerpo por cada µg de proteína total. Los anticuerpos usados para las inmunoprecipitaciones fueron un anti-BCL7A de la casa comercial Sigma-Aldrich (#HPA019762) y un anti-SMARCA4 (G7:sc-17796, Santa Cruz Biotechnology). En cada uno de los experimentos, una muestra fue incubada usando como control de uniones inespecíficas un anticuerpo irrelevante anti-conejo IgG o anti-ratón IgG para las inmunoprecipitaciones de BCL7A o SMARCA4 respectivamente.

Los anticuerpos utilizados para comprobar la co-inmunoprecipitación de ciertas proteínas mediante Western blot fueron los siguientes: anti-BCL7A monoclonal (Clone 15C, *Monoclonal Antibodies Unit*, CNIO, Madrid, España); anti-SMARCA4 policlonal (H88:sc-10768, Santa Cruz Biotechnology); anti-BRM monoclonal (D9E8B:Cat#11966, Cell Signaling); anti-SMARCC1 monoclonal (D7F8S:Cat#11956 D7F8S:Cat#11956, Cell Signaling); anti-SMARCC2 monoclonal (D809V:Cat#12760, Cell Signaling); anti-SMARCD1 monoclonal (sc-135843, Santa Cruz

Biotechnology) anti-SMARCE1 policlonal (Cat# A300-810A, Bethyl, Montgomery, TZ, EE. UU.) y anti-SMARCB1 monoclonal (D9C2:Cat# 8745, Cell Signaling).

Para el análisis mediante LC/MS-MS, los lisados de la línea celular HBL1 transducida se prepararon en buffer de lisis RIPA conteniendo inhibidores de proteasas y fosfatasa (0,2mM PMSF, 7 mM OV_4 y 1X *cOmplete Mini EDTA-free Protease Inhibitor Cocktail Tablets*). Después de cuantificar la proteína total usando el reactivo Bradford, 4mg de proteína para cada una de las condiciones se inmunoprecipitaron durante toda la noche a 4°C usando 1µg de anticuerpo anti-BCL7A (#HPA019762, Sigma-Aldrich) por cada µg de proteína total. Para cada una de las condiciones se incubó una muestra con un anticuerpo irrelevante anti-conejo IgG para analizar las uniones inespecíficas. Los inmunocomplejos se recuperaron tras añadir 200µL de *Dynabeads Protein G* (Cat#10004D, ThermoFisher Scientific) e incubar las muestras durante 3 horas a temperatura ambiente en agitación con una noria. Posteriormente, las *beads* se lavaron 3 veces consecutivas con PBS al 1X conteniendo inhibidores de proteasas. La elución final se llevó a cabo en dos pasos, añadiendo en cada paso 600µL de una solución 8M de urea en 0,1M Tris-HCl a pH8, e incubando a 26°C durante 10 minutos con agitación a 1400rpm.

4.16.1 Preparación de las muestras para el análisis proteómico

Los eluidos se digirieron mediante el protocolo estándar FASP. Brevemente, las proteínas se redujeron (TCEP 15mM, 30 minutos, temperatura ambiente), se alquilaron (CAA 50mM, 20 minutos en oscuridad, temperatura ambiente) y se digirieron secuencialmente con Lys-C (FUJIFILM Wako Chemicals, Neuss, Alemania) (relación proteína:enzima 1:50, incubación durante la noche a temperatura ambiente) y tripsina (Promega, Madison, WI, EE. UU.) (proporción proteína: enzima 1:100, 6h a 37°C). Los péptidos resultantes se desalaron usando puntas de etapa C₁₈.

4.16.2 Espectrometría de masas (Q Exactive™)

El proceso de LC-MS/MS se realizó acoplado un sistema de *HPLC UltiMate 3000* a un espectrómetro de masas *Q Exactive Plus* (ThermoFisher Scientific). Los péptidos se cargaron en una columna trampa (Acclaim™ PepMap™ 100 C₁₈ LC Columns 5 µm, 20mm de longitud) durante 3 minutos a una velocidad de flujo de 10 µL/min en 0,1% Formaldehído (FA). Luego, los péptidos se transfirieron a una columna analítica (PepMap RSLC C18 2 µm, 75 µm x 50 cm) y se separaron usando un gradiente lineal efectivo de 90 minutos (tampón A: 4% de Acetonitrilo (ACN), 0,1% de FA; tampón B: 100% de ACN, 0,1 % FA) a un caudal de 250 nL/min. El gradiente utilizado fue: 0-5 min 4% B, 5-7.5 min 6% B, 7.5-60 min 17.5% B, 60-72.5 min 21.5% B, 72.5-80 min 25% B, 80-94 min 42.5 % B, 94-100 min 98% B, 100-104.5 min 4% B, 105-110 min 0% B. El espectrómetro

de masas se hizo funcionar en un modo dependiente de datos, con un cambio automático entre los escaneos MS (350-1400 m/z) y los escaneos MS/MS y usando lo que se denomina “*Top 15 method*” (señal umbral de intensidad $\geq 3.9E4$, $z \geq 2$). Los espectros de MS se adquirieron en el Orbitrap con una resolución de 70.000 FWHM (200 m/z) y los espectros de MS/MS con una resolución de 17.500 FWHM (200 m/z). Se usó un tiempo de exclusión activa de 26,3 segundos. Los péptidos se aislaron usando una ventana de 2Th y se fragmentaron usando una disociación de colisión de alta energía (HCD) con una energía de colisión normalizada de 27. Los valores diana de los iones fueron $3E6$ para MS (tiempo de inyección máximo de 25ms) y $1E5$ para MS/MS (tiempo de inyección máximo de 90ms).

4.16.3 Análisis de los datos generados por LC-MS/MS

Los datos crudos se procesaron con MaxQuant (v 1.6.2.6a) usando la configuración estándar y usando como referencia una base de datos de proteínas humanas (UniProtKB / Swiss-Prot, 20.373 secuencias) suplementada con contaminantes. La cuantificación por LFQ (*Label-free quantification*) se realizó por coincidencia entre corridas (ventanas de coincidencia de 0,7 minutos y ventana de alineación de 20 minutos). La carbamidometilación de las cisteínas se estableció como una modificación fija, mientras que la oxidación de las metioninas y la acetilación de proteínas en el extremo amino-terminal se establecieron como modificaciones variables. La longitud mínima del péptido se fijó en 7 aminoácidos y se permitió un máximo de dos divisiones trípticas perdidas. Los resultados se filtraron a 0,01 FDR (tanto a nivel de péptido como a nivel de proteína).

Posteriormente, el archivo "proteinGroup.txt" se cargó en ProStar (v 1.12.12) para su posterior análisis estadístico. Se requirió un mínimo de dos valores válidos de LFQ por grupo para la cuantificación. Los valores ausentes se imputaron usando la configuración estándar en Prostar (es decir, SLSA para POV y Det quantile para MEC). A continuación, se realizó un análisis mediante Limma y los valores de p se ajustaron utilizando el método Benjamin-Hochberg. Solamente las proteínas con un p-valor $< 0,05$ y una proporción muestra/control con valor $\log_2 > 1,5$ se consideraron como posibles interactores. En todos los casos, el FDR resultante fue inferior al 5%.

4.17 DATOS GENERADOS EN ESTE ESTUDIO ALMACENADOS EN REPOSITORIOS

Los datos de expresión génica (archivos BAM y número de lecturas por gen) que se han discutido en este trabajo se han depositado en la base de datos *NCBI's Gene Expression Omnibus* (GEO) y son accesibles a través del número identificador GEO Series GSE149277.

En cuanto a los datos referentes a la espectrometría de masas, los datos crudos han sido depositados en el consorcio *ProteomeXchange* a través del repositorio PRIDE. El identificador de nuestra base de datos es PXD014795.

5. RESULTADOS

5.1 ESTUDIO PILOTO EN DISTINTAS NEOPLASIAS HEMATOLÓGICAS

En un primer estudio se trabajó con una colección de líneas celulares que englobaban distintas neoplasias hematológicas. Los nombres y detalles de cada una de las líneas celulares analizadas en este estudio piloto se muestran en la Tabla M1 (Sección Materiales y métodos). Este estudio piloto se realizó con el objetivo de caracterizar un modelo de enfermedad en el que *BCL7A* pudiese tener una implicación funcional y, al mismo tiempo, encontrar posibles líneas celulares que sirvan como modelos para llevar a cabo estudios funcionales sobre la actividad de *BCL7A*. Primero decidimos medir los niveles de expresión de *BCL7A*. Para ello, se extrajo ARN de cada una de las líneas celulares y se midieron los niveles de expresión de *BCL7A* mediante qPCR (Figura 8).

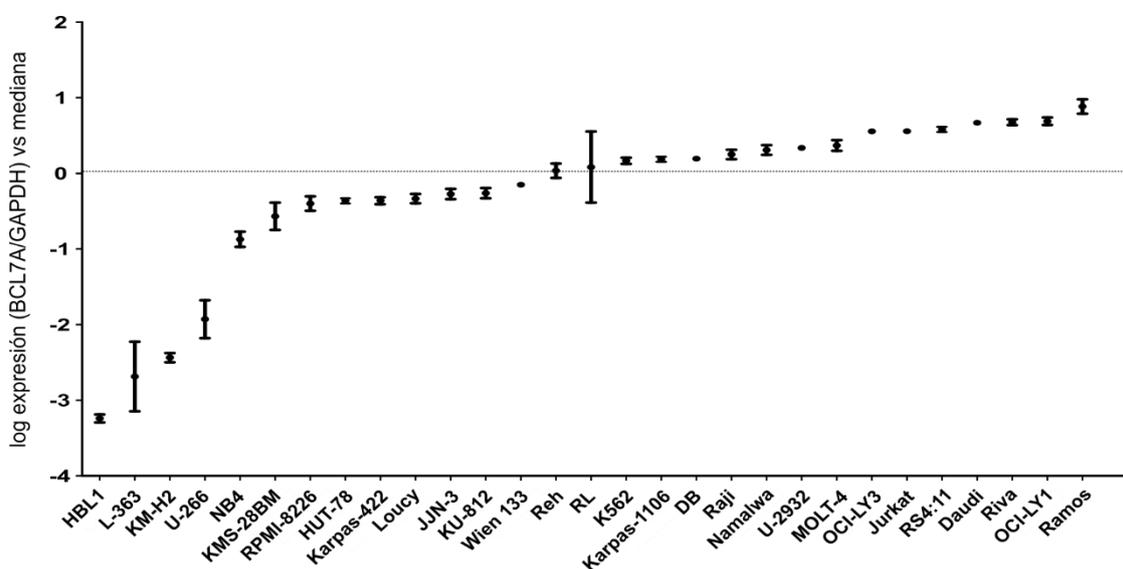


Figura 8. Análisis de expresión de *BCL7A* en una colección de 29 líneas celulares hematológicas. Análisis de expresión de *BCL7A* a nivel de ARNm mediante RT-qPCR. Los valores de expresión relativa *BCL7A/GAPDH* fueron transformados en logaritmo y se han representado en una gráfica centrada en el valor de la mediana (línea horizontal discontinua). Las barras de error representan la desviación estándar.

Basándonos en los resultados de expresión obtenidos mediante RT-qPCR, se observó que hay varias líneas celulares que tienen una expresión de *BCL7A* muy por debajo de la media. Una de las líneas celulares es HBL1, que deriva de un paciente diagnosticado con DLBCL. Otras de las líneas celulares son KM-H2 o NB4, que derivan de pacientes diagnosticados con linfoma de Hodgkin y leucemia promielocítica aguda respectivamente. Por último, otra de las líneas celulares con una expresión muy baja es L-363, que deriva de un paciente diagnosticado con mieloma múltiple y comparte origen con otras de las líneas como U-266, KMS-28BM o RPMI-8226 que también expresan bajos niveles de *BCL7A*.

El resto de líneas celulares expresan niveles más o menos homogéneos de BCL7A. Si nos fijamos en un estudio publicado anteriormente en el que analizaron la expresión de BCL7A en tejidos y células humanas (Ramos-Medina et al.), nuestros datos concuerdan con el hecho de que BCL7A se expresa en células B maduras¹⁴¹. Por tanto, es coherente detectar expresión de BCL7A en líneas derivadas de pacientes tanto de DLBCL como de linfoma de Burkitt, que representan neoplasias originadas a partir de células B. En cambio, la expresión de BCL7A es prácticamente nula en líneas derivadas de mieloma múltiple, lo cual concuerda tanto con nuestros datos como con los resultados del estudio de Ramos-Medina et al.

En cuanto a la línea celular KM-H2, se trata de una línea celular que proviene de un paciente con una forma poco común de linfoma de Hodgkin con depleción leucocitaria. Estas células presentan celularidad mixta carente de marcadores de linfocitos B como CD19 y que presentan marcadores mieloides como CD14 o CD15. Este tipo de linfoma de Hodgkin con celularidad mixta presentaba un nivel de expresión de BCL7A bajo según el estudio de Ramos-Medina et al. En nuestro estudio de expresión de BCL7A también analizamos algunas líneas celulares que derivan de pacientes con neoplasias hematológicas originadas por linfocitos T y observamos que BCL7A presentaba unos niveles de expresión similares a los observados en otras neoplasias derivadas de linfocitos B.

Durante el análisis de los datos crudos obtenidos por qPCR, se observó una anomalía en la curva de disociación en dos líneas celulares, OCI-LY1 y Karpas-1106 (**Figura 9**). Al observar la curva de disociación obtenida para la línea celular Daudi, vimos un único pico que concuerda con la amplificación de un único fragmento. La curva de disociación observada para Daudi era la misma que la observada para todas las líneas celulares a excepción de OCI-LY1 y Karpas-1106. Por el contrario, para las líneas celulares OCI-LY1 y Karpas-1106, se observó una curva de disociación con dos picos. Uno de los picos aparece a la misma temperatura de fusión obtenida en Daudi, lo que es indicativo de que podría tratarse del mismo fragmento. En cambio, el otro pico aparece a una temperatura de fusión menor, lo cual indicaría que se está amplificando un fragmento de menor tamaño. En conjunto, estos patrones observados podrían ser el resultado de una amplificación aberrante de los transcritos del gen *BCL7A*, debido a una posible anomalía genética en las líneas celulares OCI-LY1 y Karpas-1106.

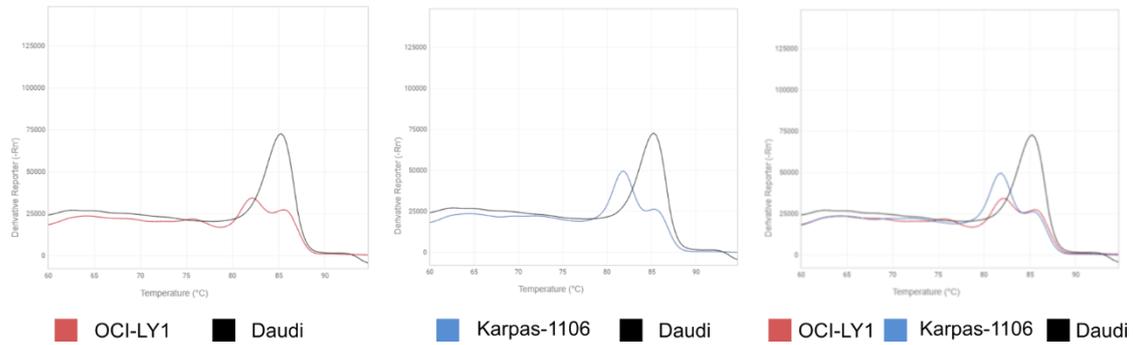


Figura 9. Curvas de disociación aberrantes obtenidas mediante RT-qPCR en líneas celulares de DLBCL. Se muestran las curvas de disociación obtenidas mediante qPCR para la línea celular OCI-LY1 (izquierda), Karpas-1106 (centro) y la superposición de ambas (derecha). La curva de disociación obtenida para la línea celular Daudi se muestra en cada uno de los gráficos como referencia de la amplificación de un transcrito *BCL7A* silvestre.

Además del análisis de expresión de *BCL7A* por RT-qPCR en el panel de líneas celulares seleccionado, decidimos realizar un análisis en busca de mutaciones en el gen *BCL7A*. Inicialmente se diseñó una estrategia para secuenciar *BCL7A* en las distintas líneas celulares a nivel de ARN mensajero (ARNm). Dado que no contábamos con técnicas de secuenciación masiva, secuenciar el ARN suponía el abordaje más eficiente en cuanto a número de PCRs por línea celular. En la **figura 10** se muestra un esquema del transcrito de *BCL7A*, el cual está compuesto de 6 exones, que dan lugar a la expresión de dos isoformas debido a un *splicing* alternativo en el exón 5. Se utilizaron dos parejas de cebadores para amplificar dos regiones distintas que de forma conjunta engloban la secuencia codificante completa de *BCL7A*.

```

TGTGAGTGTG TGCGTGTGAG AGTCCGAGTC TCTGTCCCG AGTGAGTGAG
CGGCGGGCGG GCGCGAGTGT GGCCGCGCGG GAGCGCGAGC AGGACCCGGC
GGGCGCGCTC CCCAGCCTCC GTCTCCCGC CGAACCATG TCGGGCAGGT
CGGTTCGAGC CGAGACGAGG AGCCGGGCCA AAGATGATAT CAAGAGGGTC
ATGGCGGCGA TCGAGAAAGT GCGCAAATGG GAGAAGAAAT GGGTGACCGT
TGGTGACACA TCCCTACGAA TCTACAAATG GGTCCCTGTG ACGGAGCCCA
AGGTGATGA CAAAAACAAG AATAAGAAA AAGCAAGGA CGAGAAGTGT
GGCTCAGAGG TGACCACTCC GGAGAACAGT TCCTCCCCAG GGATGATGGA
CATGCATGAC GATAACAGCA ACCAGAGTC CATCGCAGAT GCCTCCCCCA
TCAAACAGGA GAACAGCAGC AACTCCAGCC CCGCTCCAGA CCCCAACTCC
ACTGTGCCCCA GGGACGGCAC CGAGGCCAAG BTGGATGAGG CCCAGGCTGA
TGGGAAGGAG CACCCAGGAG CTGAAGATGC TTCTGATGAG CAGAATTCAC
AGTCTCGAT GGAACATTCG ATGAACAGCT CAGAGAAAGT AGATCGGCAG
CCGTCTGGAG ACTCGGGTCT GGCCGCAGAG ACGTCTGCAA TCTCTCAGGT
ACCTCGCTCG AGGTCTCAGA GGGGCAGCCA GATCGGCCCG GAGCCCATTG
GGTTGTCGGG GGATCTGGAA GGAGTGCCAC CCTCTAAAA GATGAACTG
GAGGCTCTC AACAAAACTC CGAAGAGATG TAGACGATGC TTTAAAGCCT
CCGATCCATG TTCCATGGAA GGTACATCAG CAATTAATTC TAGAGCAACT
TTGCCCCAGC GATTCTCTT GGGTCCGAAC AGAACTACTA ACGTTTCAAG

```

Figura 10. Representación de la secuencia de *BCL7A* a nivel de ADNc desde la posición 1 a la posición 1000. La secuencia codificante se muestra en azul, mientras que las regiones no traducidas (UTRs) se muestran en rojo. Las letras negras representan el primer y último nucleótido de cada exón. Las regiones coloreadas en verde o amarillo representan las dos parejas de cebadores (directo y reverso) utilizados para amplificar dos regiones independientes del transcrito de *BCL7A*. Finalmente, en color naranja se muestra la región de 63pb que está presente en la isoforma larga pero no en la isoforma corta de *BCL7A* debido al *splicing* alternativo del exón 5.

En la **figura 11A** podemos ver un ejemplo del resultado de amplificación obtenido para algunas de las líneas celulares del estudio piloto. Si nos fijamos únicamente en una de las líneas celulares como Ramos, vimos que la amplificación utilizando la primera pareja de cebadores (1; color verde) dio lugar a un único fragmento de 478pb. En cambio, al utilizar la segunda pareja de cebadores (2; color amarillo) se obtuvo un patrón de doble banda. Esto es debido a que esta región contiene la parte diferencial que determina las distintas isoformas de BCL7A. Por lo tanto, se amplificaron dos bandas de 420 y 483pb que se corresponden a las isoformas corta y larga respectivamente.

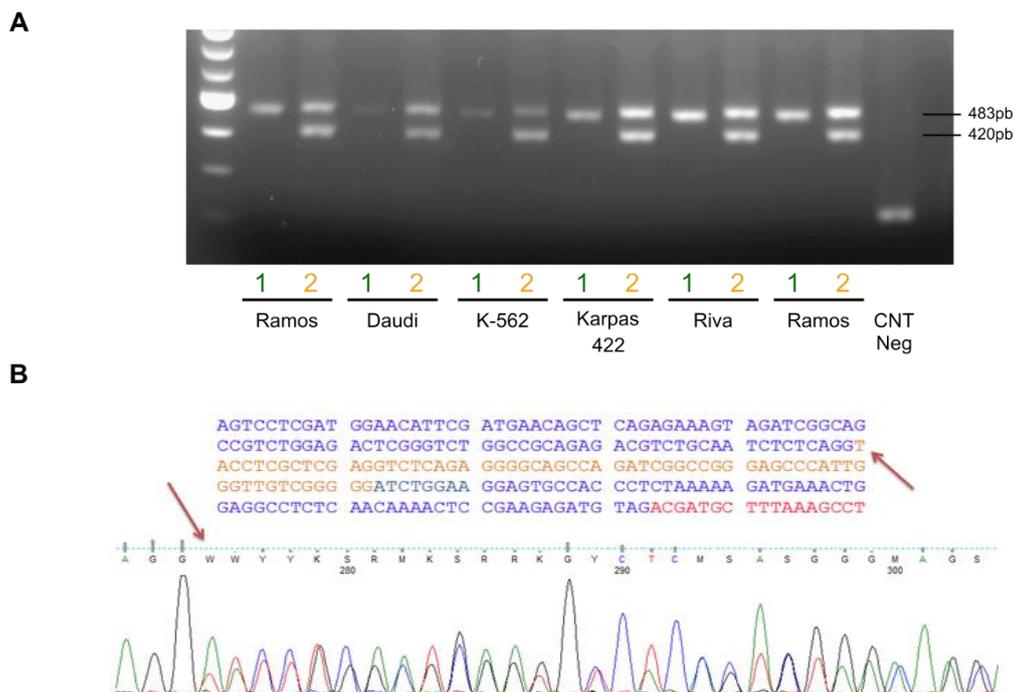


Figura 11. Amplificación de BCL7A mediante PCR para secuenciación de Sanger. A) Se amplificó BCL7A a partir de ADNc en diferentes líneas celulares. Se amplificaron de forma independiente dos regiones de BCL7A con las dos parejas de cebadores 1 y 2 que se muestran en la figura 10 de color verde y amarillo respectivamente. Se utilizó un control negativo (CNT Neg) que no contenía molde para la PCR. **B)** Visualización del cromatograma obtenido por secuenciación de Sanger de la banda de 483pb obtenida al amplificar la región 2 de BCL7A. Las flechas rojas indican la posición del nucleótido donde empieza la región diferencial de las isoformas de BCL7A. Debido a la doble banda no se pudo obtener una lectura correcta.

Al secuenciar la banda de 483pb, obtenida al amplificar la segunda región de BCL7A, observamos que se obtenía una lectura que impedía un correcto análisis (**Figura 11B**). Por lo tanto, con el objetivo de mejorar el análisis, diseñamos una nueva estrategia de amplificación. Lo que hicimos fue amplificar la secuencia codificante completa de BCL7A utilizando cebadores flanqueantes a la misma. De esta forma obtuvimos dos bandas de tamaños 796 y 859pb relativas a la isoforma corta y la isoforma larga de BCL7A respectivamente (**Figura 12**). Posteriormente, escindimos y purificamos únicamente la banda de 859pb. Para su secuenciación utilizamos cebadores directos

y reversos que se unían a la región de 63pb propia de la isoforma larga y de esa manera conseguimos obtener secuenciaciones que cubrían la secuencia codificante completa de *BCL7A* para muchas de las líneas celulares del estudio piloto. Las líneas celulares para las que no se pudo obtener amplificación fueron las que tenían niveles de expresión igual o menor a la línea celular RPMI-8226 (Ver Figura 8).

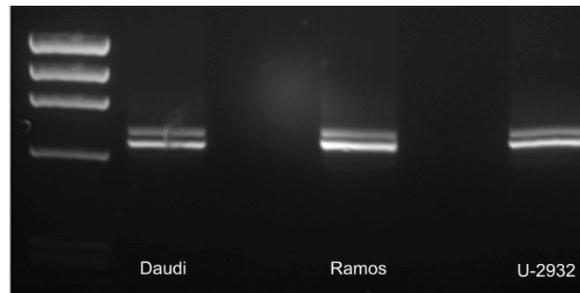


Figura 12. Amplificación de la región codificante completa de *BCL7A* a nivel de ADNc. Se observa un patrón de doble banda debido a la amplificación tanto de la isoforma corta (796pb) como de la isoforma larga (859pb) de *BCL7A* en las tres líneas celulares indicadas.

Tras analizar las secuenciaciones para las todas las líneas celulares del estudio piloto, se encontraron mutaciones en dos de ellas: DB y OCI-LY1. Ambas líneas celulares derivan de pacientes diagnosticados con DLBCL. En este punto, no vamos a definir y mostrar las mutaciones encontradas ya que se discutirán en la próxima sección.

La conclusión del estudio piloto es que las líneas celulares de DLBCL parecían aglomerar distintas anomalías observadas al amplificar *BCL7A*. Primero, la ausencia de expresión en la línea celular HBL1. Por otro lado, se encontraron mutaciones en líneas como DB y OCI-LY1. Finalmente, también se observó la presencia de posibles transcritos aberrantes en las líneas celulares OCI-LY1 o Karpas-1106 en base a las curvas de disociación obtenidas mediante qPCR. En líneas derivadas de pacientes con mieloma múltiple la expresión de *BCL7A* es muy baja, lo cual concuerda con la baja expresión fisiológica que muestran las células plasmáticas¹⁴¹. En otras líneas como KM-H2 o NB4 también hemos encontrado ausencia de expresión, pero representan patologías muy poco comunes y no disponíamos de líneas celulares adicionales para corroborar los resultados obtenidos.

5.2 ESTUDIO MUTACIONAL EN DLBCL

Tras el cribado realizado en distintas líneas celulares de origen hematológico, pudimos comprobar que las mutaciones u otras anomalías se encontraban en líneas celulares que tenían como denominador común neoplasias en células B y más concretamente en líneas celulares derivadas de pacientes con DLBCL. Este tipo de linfoma tiene la característica de que se origina

a partir de células B que han quedado ancladas en algún momento de su diferenciación a través del CG. Teniendo en cuenta que la expresión de *BCL7A* se modula en función de los diferentes estadios de diferenciación de una célula B¹⁴¹, es concebible que *BCL7A* pueda tener importantes implicaciones tanto en la fisiología del CG como en la fisiopatología del DLBCL.

Con el objetivo de ampliar nuestro análisis a un mayor número de líneas celulares derivadas de DLBCL, conseguimos aglomerar un panel con un total de 41 líneas celulares derivadas de pacientes diagnosticados con DLBCL, así como 38 muestras de pacientes que habían desarrollado DLBCL. También, se incluyó un panel de 15 muestras de pacientes diagnosticados con FL, un tipo de linfoma no Hodgkin que comparte grandes similitudes con el DLBCL y que en determinadas ocasiones evoluciona histológica y molecularmente a DLBCL¹⁵².

Con el fin de analizar el estado mutacional de *BCL7A* en DLBCL, se llevó a cabo la secuenciación, en este caso a nivel de ADN, de cada uno de los exones del gen en cada una de las muestras previamente mencionadas (total n=94 muestras incluyendo líneas celulares y pacientes). En el **apéndice III** se resumen las mutaciones encontradas tanto en líneas celulares como en muestras de pacientes. A continuación, se detallan las mutaciones y otras anomalías encontradas a raíz del estudio mutacional.

5.2.1 Variantes genéticas en *BCL7A* en líneas celulares

De las 41 líneas celulares incluidas en el análisis, siete de ellas presentaron alguna variación genética en *BCL7A*, lo que representa el 17%. Se identificaron principalmente dos tipos de mutaciones: mutaciones con cambio de sentido y mutaciones que afectan al sitio de *splicing*. A continuación, se muestran en detalle las variaciones genéticas encontradas, desglosadas por tipo de mutación.

5.2.1.1 Mutaciones con cambio de sentido

Las líneas celulares DB, OCI-LY7 y VAL presentaron una mutación con cambio de sentido en heterocigosis (**Figura 13**). En las tres líneas celulares, las mutaciones se localizan en el exón 1, aunque en distintas posiciones (**Apéndice III**). En la línea celular DB se detectó la mutación c.33G>C, la cual da lugar a un cambio de arginina por serina en el codón 11 (p.R11S). En cuanto a OCI-LY7, esta línea celular presenta la mutación c.85C>T, que genera un cambio de aminoácido en el codón 29 de la secuencia peptídica (p.R29C). En el caso de VAL, la mutación se encuentra en la posición 32 de la secuencia codificante (c.32G>C). A nivel de proteína, se produce el cambio del aminoácido arginina por treonina, también en el codón 11 (p.R11T). En los tres casos mencionados, la mutación implica un cambio del aminoácido básico arginina por un aminoácido

polar sin carga; sin embargo, se desconoce la consecuencia funcional que pueden tener estas alteraciones en la función de la proteína.

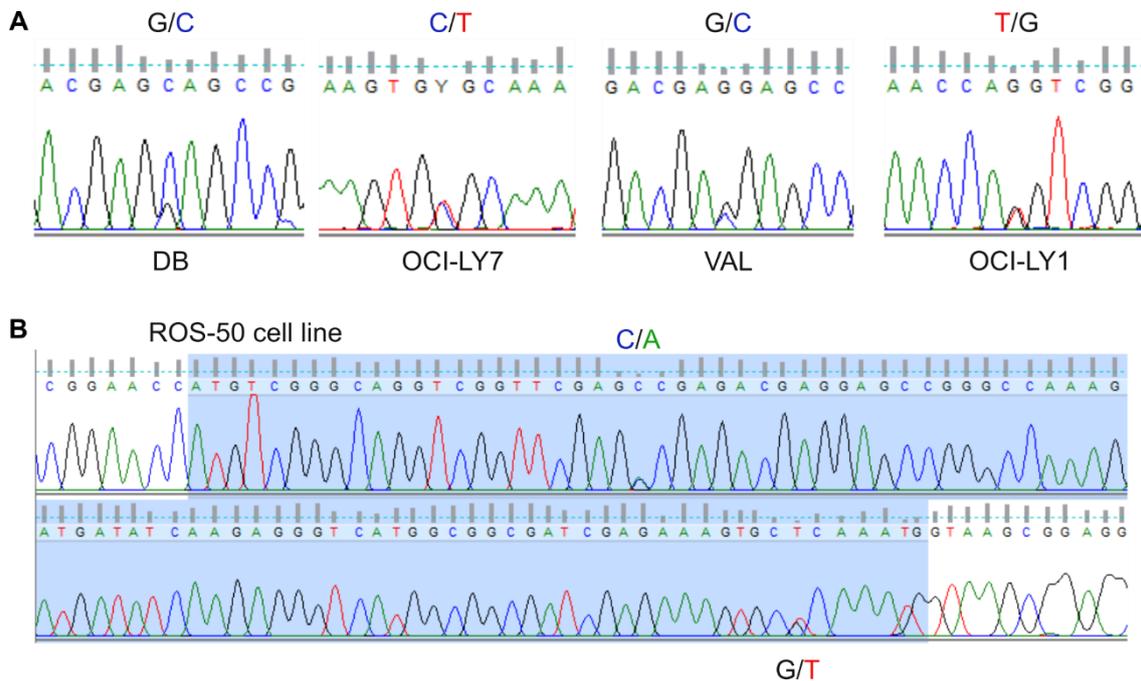


Figura 13. Mutaciones con cambio de sentido y mutación puntual en el codón de inicio en líneas celulares de DLBCL. A) Cromatogramas representativos, obtenidos mediante secuenciación de Sanger, que muestran tanto las mutaciones con cambio de sentido encontradas en las líneas celulares DB, OCI-LY7 y VAL, como la mutación puntual en el codón de inicio encontrada en la línea celular OCI-LY1. **B)** Cromatograma completo del exón 1 en la línea celular ROS-50 donde se muestran las dos mutaciones heterocigotas obtenidas para esta línea celular.

En el caso de la línea celular ROS-50 se identificaron dos mutaciones en el exón 1. Ambas son mutaciones con cambio de sentido y se presentan en heterocigosis (**Figura 13B**). La primera de ellas se localizó en la posición 23 de la secuencia codificante (c.23C>A) y provoca un cambio de alanina por ácido aspártico (p.A8D). La segunda mutación se encontró en posición 86 (c.86G>T) y también origina un cambio de aminoácido en la secuencia peptídica, en este caso de arginina por leucina en el codón 29 (p.R29L). Mediante la técnica conocida como “TA cloning”, que permite la clonación y posterior secuenciación individualizada de cada uno de los alelos a través de un vector comercial, pudimos determinar que las dos mutaciones heterocigotas presentes en la línea celular ROS-50 se localizan en alelos diferentes (resultados de secuenciación no mostrados).

5.2.1.2 Mutación puntual en el codón de inicio de la traducción

En el caso concreto de la línea celular OCI-LY1, se encontró una mutación heterocigota en el codón de inicio de la traducción (c.2T>G). Esta mutación, a diferencia del resto de mutaciones con cambio de sentido mencionadas, no genera un cambio de aminoácido, sino que implica que

la maquinaria de traducción no pueda reconocer el inicio de la secuencia codificante del gen y, por tanto, se genera una pérdida de expresión de proteína a partir del alelo mutado.

5.2.1.3 Mutaciones de *splicing*

En las líneas celulares Karpas-1106, OCI-LY1 y VAL se encontró una mutación en heterocigosis en la misma posición, el primer nucleótido del intrón 1 (**Figura 14**). Esta posición, ocupada por el nucleótido guanina (G), es muy importante para la maquinaria de procesamiento postranscripcional ya que representa el sitio donador de *splicing*¹⁵³. Como desglosaremos en las próximas secciones, esta mutación conlleva importantes implicaciones funcionales.

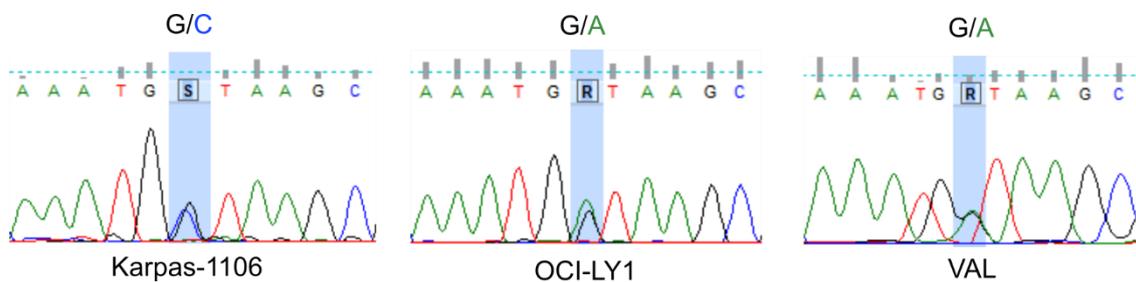


Figura 14. Mutaciones de *splicing* en líneas celulares de DLBCL. Cromatogramas obtenidos mediante secuenciación de Sanger que muestran las mutaciones heterocigotas en el sitio donador de *splicing* del intrón 1 en las líneas celulares Karpas-1106, OCI-LY1 y VAL.

5.2.1.4 Deleción en *HBL1*

Un caso particular fue el de la línea celular HBL1. En esta línea celular no se consiguió amplificar el exón 1 de *BCL7A* a nivel de ADN después de varios intentos (**Figura 15A**). En cambio, sí que se amplificaron correctamente el resto de exones de *BCL7A* como se puede ver para los exones 2 y 3 en la **figura 15B**. Este resultado sugiere que HBL1 podría presentar una deleción o translocación que afecte al exón 1.

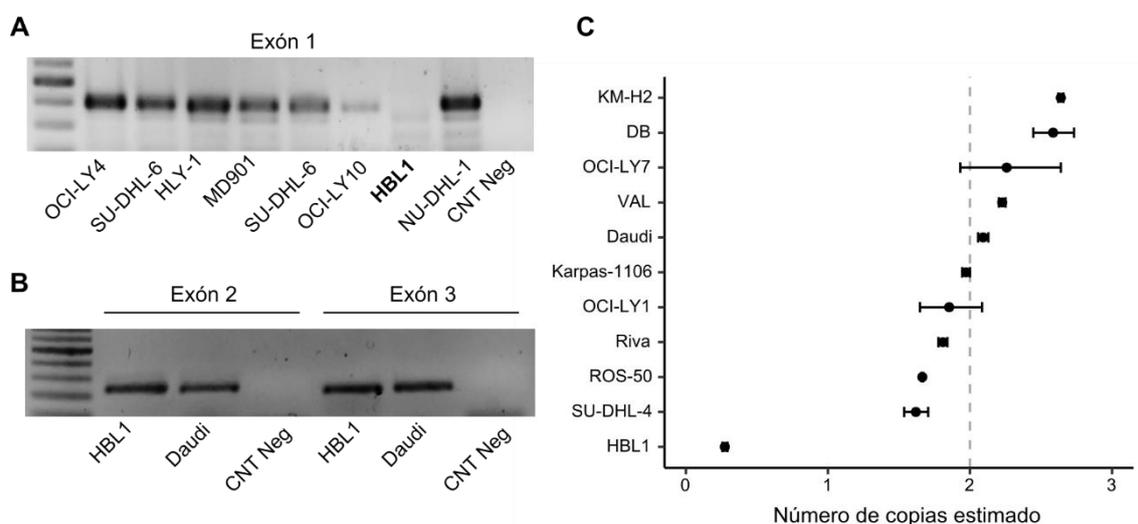


Figura 15. La línea celular HBL1 presenta una delección del locus del exón 1 de *BCL7A*. **A)** Electroforesis en geles de agarosa de los productos de PCR obtenidos al amplificar el exón 1 de *BCL7A* en distintas líneas celulares de DLBCL. Se utilizó un control negativo (CNT Neg) que no contenía molde para la PCR. **B)** Electroforesis en geles de agarosa de los productos de PCR obtenidos al amplificar los exones 2 y 3 de *BCL7A* en las líneas celulares HBL1 y Daudi. Se utilizó un CNT Neg que no contenía molde para la PCR. **C)** Análisis de variación de número de copias (sondas TaqMan) para *BCL7A* en las líneas celulares mostradas. Los valores de $\Delta\Delta Ct$, calculados por el software *CopyCaller*, se promediaron entre las dos repeticiones del ensayo. Los valores del número de copia se estimaron a partir del promedio de los valores de $\Delta\Delta Ct$. Las barras de error representan el rango. La línea discontinua vertical representa el número de copias teórico para un genoma diploide.

Para confirmar la posibilidad de que HBL1 presente una delección se llevó a cabo un análisis de variación de número de copias utilizando distintas líneas celulares entre las que se encontraba HBL1. Como se puede apreciar en la **figura 15C**, prácticamente todas las líneas presentan un número de copias en torno a 2, indicativo de la presencia de dos alelos. Los valores de $n=2$ para Karpas-1106, OCI-LY1, ROS-50 o VAL concuerdan con los datos previamente mostrados (**Figura 14**). En cambio, para HBL1 el número de copias es cercano a cero. Esto hace pensar que, efectivamente, hay una delección de la región cromosómica en torno al exón 1 de *BCL7A*, o bien podría haber una translocación cromosómica. El análisis de variación de número de copias también nos indicó un valor cercano a $n=3$ para las líneas celulares DB y KM-H2. Este dato concuerda con el hecho de que ambas líneas celulares han sido caracterizadas como triploides ($n=3$; información recogida en la base de datos *DSMZ*). Además, el dato de $n=3$ concuerda con nuestros datos de secuenciación a nivel de ADN para la línea celular DB (**Figura 13A**), donde el cromatograma muestra que el pico relativo al nucleótido mutante citosina (C) tiene el doble de tamaño que el nucleótido silvestre guanina (G). Esto es indicativo de una proporción alélica mutante:silvestre de 2:1. Por tanto, la línea celular DB expresaría dos copias del alelo mutante R11S-*BCL7A* y una copia del alelo silvestre.

5.2.2 Variantes genéticas en *BCL7A* en muestras de pacientes

De las 53 muestras de pacientes incluidas en el análisis (38 DLBCL + 15 FL), 3 de ellas presentaron una mutación en *BCL7A*. Se identificaron 2 mutaciones heterocigotas con cambio de sentido en 2 pacientes con DLBCL (prevalencia del 5,26%) (**Apéndice II**). Las mutaciones encontradas fueron c.32G>T en el paciente D29, que da lugar a un cambio de arginina por metionina en el codón 11 (p.R11M), y la mutación c.17T>C en el paciente D32, que da lugar a una sustitución de valina por alanina en el codón 6 (p.V6A). Ambas mutaciones se localizan en el exón 1 del gen. Por otro lado, en un paciente diagnosticado con FL se identificó una mutación en heterocigosis en el primer sitio donador de *splicing*, representando una prevalencia mutacional del 6,6 % en FL.

5.2.3 SNPs en regiones codificantes de *BCL7A*

Otro tipo de variantes genéticas que pueden darse en la secuencia de ADN son los polimorfismos de nucleótido único (SNP). Este tipo de variaciones deben darse en al menos un 1% de la población para ser consideradas como SNPs. Tanto en los pacientes como en las líneas celulares analizadas en este estudio se detectaron tres SNPs en *BCL7A*, todos ellos descritos previamente en bases de datos. En la **tabla 2** se resumen los SNPs encontrados y se detalla su localización en el gen, su frecuencia genotípica y la frecuencia alélica en una población control dentro del proyecto *1000 Genomes*.

El SNP localizado en el exón 2 corresponde a la variante c.174C>T, presente en 10 muestras en heterocigosis y en 2 muestras en homocigosis. Se trata de un SNP sinónimo, es decir, una mutación que no genera un cambio de aminoácido en la proteína, por lo que probablemente no tenga repercusión funcional. El SNP identificado en el exón 4 corresponde a la variante c.359A>C, que origina un cambio de aminoácido en el codón 120 (N120T). Este SNP no sinónimo, que implica un cambio aminoácido, se encontró en 8 muestras (8,42%), y se observa en heterocigosis en el 50% de esos casos. Por último, el análisis del exón 6 mostró otro SNP sinónimo, L210L, en más de la mitad de los pacientes y líneas celulares (72,34%), presente tanto en heterocigosis como en homocigosis. La frecuencia alélica de este SNP descrita en los diferentes grupos poblacionales del proyecto *1000 Genomes* es también muy alta (**Tabla 2**), por lo que se trata de una variación genética muy común en la población y, probablemente, no tiene consecuencias en la funcionalidad de *BCL7A*.

Tabla 2. SNPs detectados en *BCL7A* mediante secuenciación de Sanger en líneas celulares y pacientes

Referencia del SNP*	Variación de nucleótido	Localización en el gen	Variación de aminoácido	Frecuencia del SNP (n=94)	Frecuencia genotípica (n=94)	MAF** (1000 Genomes)
rs2272132	c.174C>T	Exón 2	D58D	12 (12,77%)	CC	83 (87,23%)
					TC	10 (10,64%)
					TT/T-	2 (2,13%)
rs34821485	c.359A>C	Exón 4	N120T	8 (8,51%)	AA	86 (91,49%)
					AC	4 (4,25%)
					CC/C-	4 (4,25%)
rs28629903	c.628T>C	Exón 6	L210L	68 (72,34%)	TT	26 (27,66%)
					TC	29 (30,85%)
					CC/C-	39 (41,49%)

* Referencia de la base de datos dbSNP.

** MAF representa la frecuencia de aparición del alelo menos común en la población. En este caso, los grupos poblacionales incluidos en el proyecto *1000 Genomes*. Los datos de MAF se obtuvieron de la base de datos dbSNP (<https://www.ncbi.nlm.nih.gov/SNP>).

*** En este caso el alelo de menor frecuencia encontrado en nuestras muestras es el considerado silvestre (timina)

5.2.4 Variantes genéticas en regiones no codificantes

Además de las mutaciones encontradas en la secuencia codificante de *BCL7A*, también se detectaron mutaciones en regiones no codificantes del gen. Para llevar a cabo la amplificación de todos los exones de *BCL7A* se habían utilizado cebadores que anillaban en las regiones intrónicas flanqueantes. Por tanto, también se pudo secuenciar una pequeña parte de estas regiones. En concreto, para la amplificación del exón 1 se utilizó un cebador directo que anillaba en la región 5'-UTR y un cebador reverso en el primer intrón. Los resultados de secuenciación mostraron mutaciones puntuales e inserciones en diferentes posiciones de la región 5'UTR en la mayoría de pacientes y líneas celulares. La información de estas mutaciones se resume en la **tabla 3**.

Tabla 3. Variantes genéticas en regiones no codificantes

Referencia de la variación genética	Variación nucleótido	Localización en el gen	MAF (1000 genomes)
rs776212842	c.-134_-133insAG	5'UTR	Desconocida
rs1678962	c.-135T>A	5'UTR	A = 27,28%
rs7298159	c.-162A>G	5'UTR	Desconocida
rs182018224	c.-164G>A	5'UTR	Desconocida
rs576752089	c.-165_-164insAT	5'UTR	Desconocida
No descrita	c.-166G>A	5'UTR	-
No descrita	c.-176T>G	5'UTR	-
rs2280574	c.624+83T>C	Intrón 5	C=40,43%

También se identificó una mutación en el intrón 5 en el 25% de las muestras, aunque en una posición alejada del sitio de *splicing*, por lo que no se considera que pudiese tener efecto en el procesamiento del ARNm.

5.3 CARACTERIZACIÓN MOLECULAR DE LOS TRANSCRITOS DERIVADOS DE LAS MUTACIONES DE SPLICING

Según el análisis mutacional llevado a cabo en muestras de DLBCL y FL, las mutaciones de *splicing* representaban un 33,3% del total de las 12 mutaciones encontradas (**Apéndice II**), solamente teniendo en cuenta mutaciones en la secuencia codificante o sitios de *splicing*. La importancia de esta mutación reside en que se localiza en el sitio donador de *splicing*, una zona en la unión exón-intrón que es reconocida por el espliceosoma. Este complejo está implicado en el

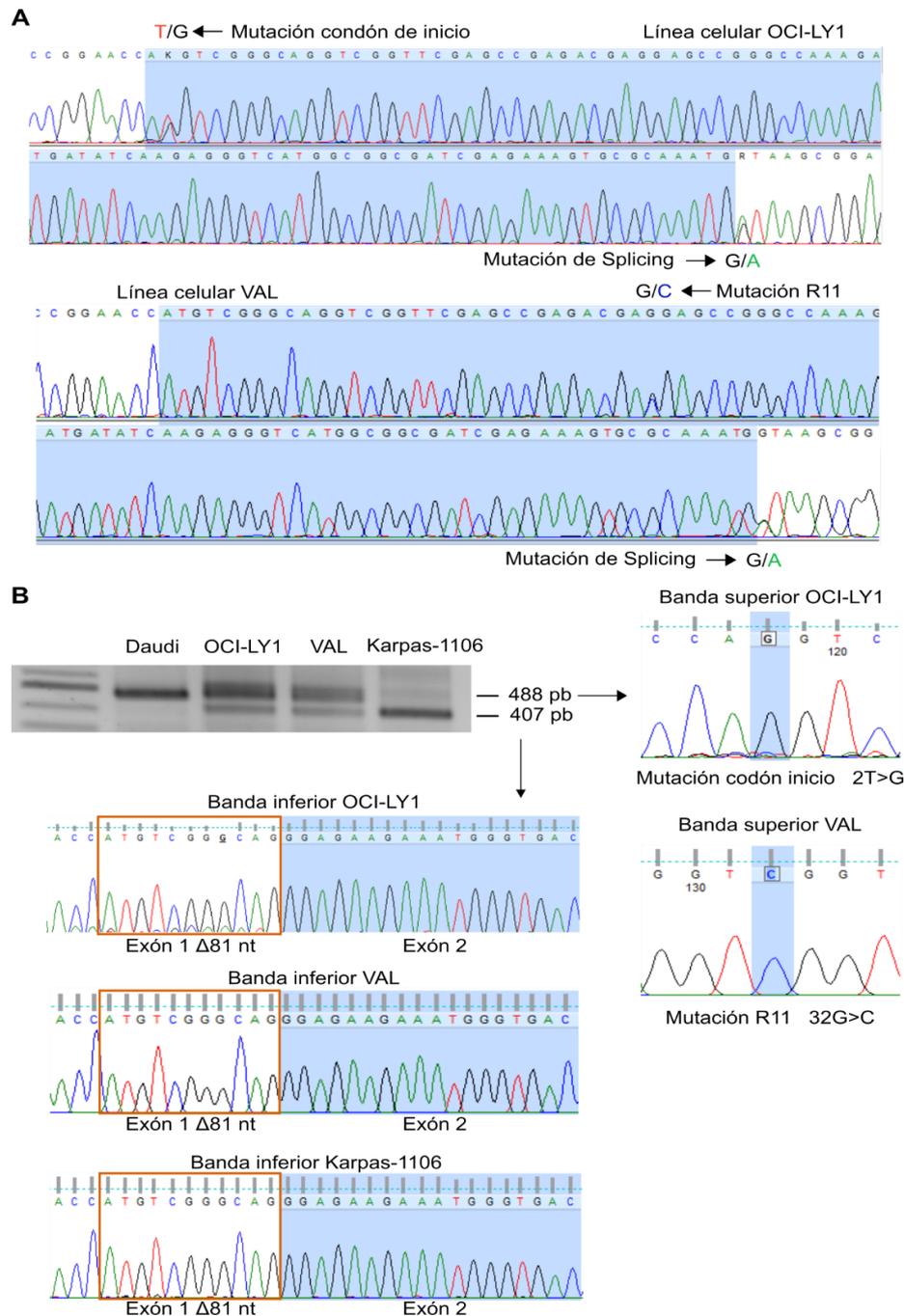


Figura 16. Diferentes mutaciones impiden la expresión de la forma silvestre de BCL7A en algunas líneas celulares de DLBCL. A) Cromatograma obtenido mediante secuenciación de Sanger del exón 1, a nivel de ADN, en las líneas celulares OCI-LY1 (panel superior) y VAL (panel inferior). En ambas líneas celulares se indican las mutaciones heterocigotas caracterizadas. La secuencia codificante del exón 1 se muestra resaltada en color azul. **B)** Secuenciación de Sanger de los diferentes productos de PCR obtenidos, a nivel de ADNc, en las líneas celulares OCI-LY1, VAL y Karpas-1106. El diseño de la PCR se muestra en la Figura 17C. En cuanto a la línea celular OCI-LY1, la banda superior tiene una mutación puntual en el codón de inicio de la traducción, mientras que la banda inferior tiene un codón de inicio ATG silvestre, pero es 81nt más corta debido a la mutación de *splicing*. En el caso de la línea celular VAL, la banda superior tiene un codón R11 mutado mientras que la banda inferior es 81nt más corta debido a la mutación de *splicing*. Por último, la banda que se obtiene en Karpas-1106 es equivalente a la obtenida en las otras dos líneas celulares y es 81nt más corta que el amplicón de tamaño silvestre, el cual se muestra para la línea celular Daudi, utilizada como control.

procesamiento del transcrito primario de ARNm (o pre-ARNm) mediante la eliminación de intrones para dar lugar al ARNm maduro. Por tanto, una mutación en esta región podría alterar el patrón normal de corte y empalme (*splicing*) y generar transcritos maduros aberrantes.

Las tres líneas celulares en las que se detectaron mutaciones de *splicing* son Karpas-1106, OCI-LY1 y VAL (**Figura 14**). Mientras que para Karpas-1106 solamente se observó la mutación de *splicing* como única mutación en la secuencia codificante, para las líneas celulares OCI-LY1 y VAL se detectaron mutaciones adicionales en el codón de inicio y en R11 respectivamente, tal y como podemos ver en las secuenciaciones de Sanger completas del exón 1 (**Figura 16A**). Para determinar qué consecuencias tiene dicha mutación en el ARNm maduro, se procedió a amplificar una región de BCL7A a nivel de ADNc. Para ello, se utilizó un cebador directo que anillaba en el exón 1, concretamente en la región 5'-UTR, y un cebador reverso en el exón 4. Los resultados de la amplificación se muestran en la **figura 16B**. El patrón de bandas obtenido en OCI-LY1, Karpas-1106 y VAL se comparó con el patrón observado en Daudi, línea celular que expresa la forma silvestre de BCL7A y que se utilizó como control de la amplificación. En el caso de las tres líneas celulares que albergan las mutaciones de *splicing* se observó una banda de menor tamaño con respecto al observado en Daudi. Las secuenciaciones de los amplicones de menor tamaño revelaron que existe una delección de 81nt a nivel de ARNm (**Figura 16B**).

La conclusión que se puede extraer de estos resultados es que, al estar mutado el sitio donador de *splicing*, el espliceosoma reconoce como nueva secuencia donadora de *splicing* una secuencia GT que se encuentra dentro del propio exón 1 (**Figura 17A**). Este fenómeno se conoce como generación de un *splicing* críptico y en el caso de BCL7A, se genera un ARNm maduro que es 81nt más corto (**Figura 17B-C**). A pesar del *splicing* críptico, el marco de lectura permanece intacto y, como veremos más adelante, la proteína resultante pierde 27 aminoácidos (S5_W31del) en la región amino-terminal, dando lugar a un BCL7A mutante que nosotros hemos nombrado $\Delta 27$ -BCL7A.

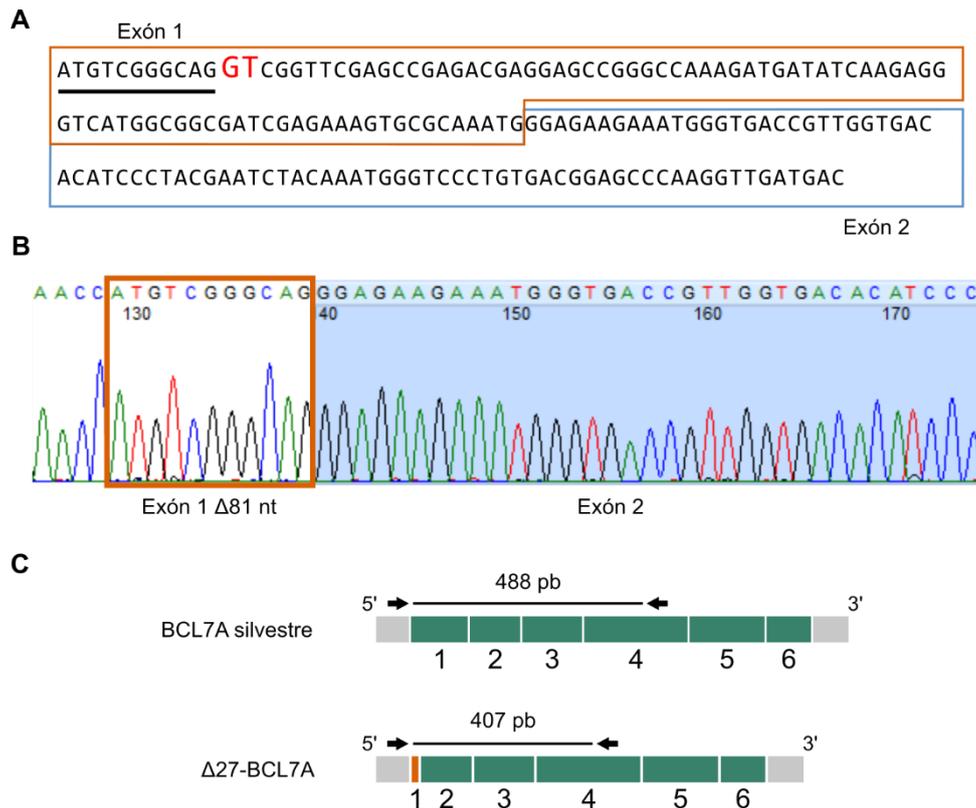


Figura 17. Las mutaciones de *splicing* generan un BCL7A mutante que pierde gran parte del exón 1. **A)** Representación esquemática del sitio de *splicing* crítico (rojo) que la maquinaria de corte y empalme reconoce cuando el sitio de *splicing* original está mutado. Los nucleótidos subrayados en el exón 1 se mantienen presentes en el ARNm de la forma mutante Δ 27-BCL7A, mientras que el resto de nucleótidos del exón 1 se pierden durante el procesamiento del transcrito aberrante. **B)** Secuenciación de Sanger del ARNm de Δ 27-BCL7A en la línea celular Karpas-1106. Se trata de un ARNm más corto que carece de 81nt del exón 1 debido a la mutación en el sitio *splicing* (92+1G>C). Este ARNm mutante mantiene el marco de lectura de BCL7A. **C)** Representación gráfica de las isoformas cortas de BCL7A tanto para un BCL7A silvestre como para un BCL7A mutante (Δ 27-BCL7A). El exón 1 de la forma mutante quedaría reducido a unos pocos nucleótidos (señalados en color naranja). Debido a los diferentes tamaños de los transcritos, obtenemos amplicones de diferentes tamaños mediante amplificación por PCR de la región de BCL7A indicada: un producto de PCR de 488pb al amplificar BCL7A silvestre; otro de 407pb al amplificar el transcrito mutante Δ 27-BCL7A. Las flechas indican la dirección: cebador directo (\rightarrow) o reverso (\leftarrow) y la ubicación de los mismos: 5'-UTR (directo) y Exón 4 (reverso).

En el caso de la línea celular OCI-LY1, la secuenciación de la banda superior (**Figura 16B**), del mismo tamaño que la banda silvestre, confirmó que la mutación puntual en el codón de inicio está presente en el alelo contrario al que presenta la mutación de *splicing*. También se secuenció la banda superior obtenida para la línea celular VAL encontramos la mutación definida en el codón R11 (**Figura 16B**). En conclusión, tanto en OCI-LY1 como en VAL, las mutaciones de *splicing* y las mutaciones que implican una segunda variación genética se encuentran en distintos alelos.

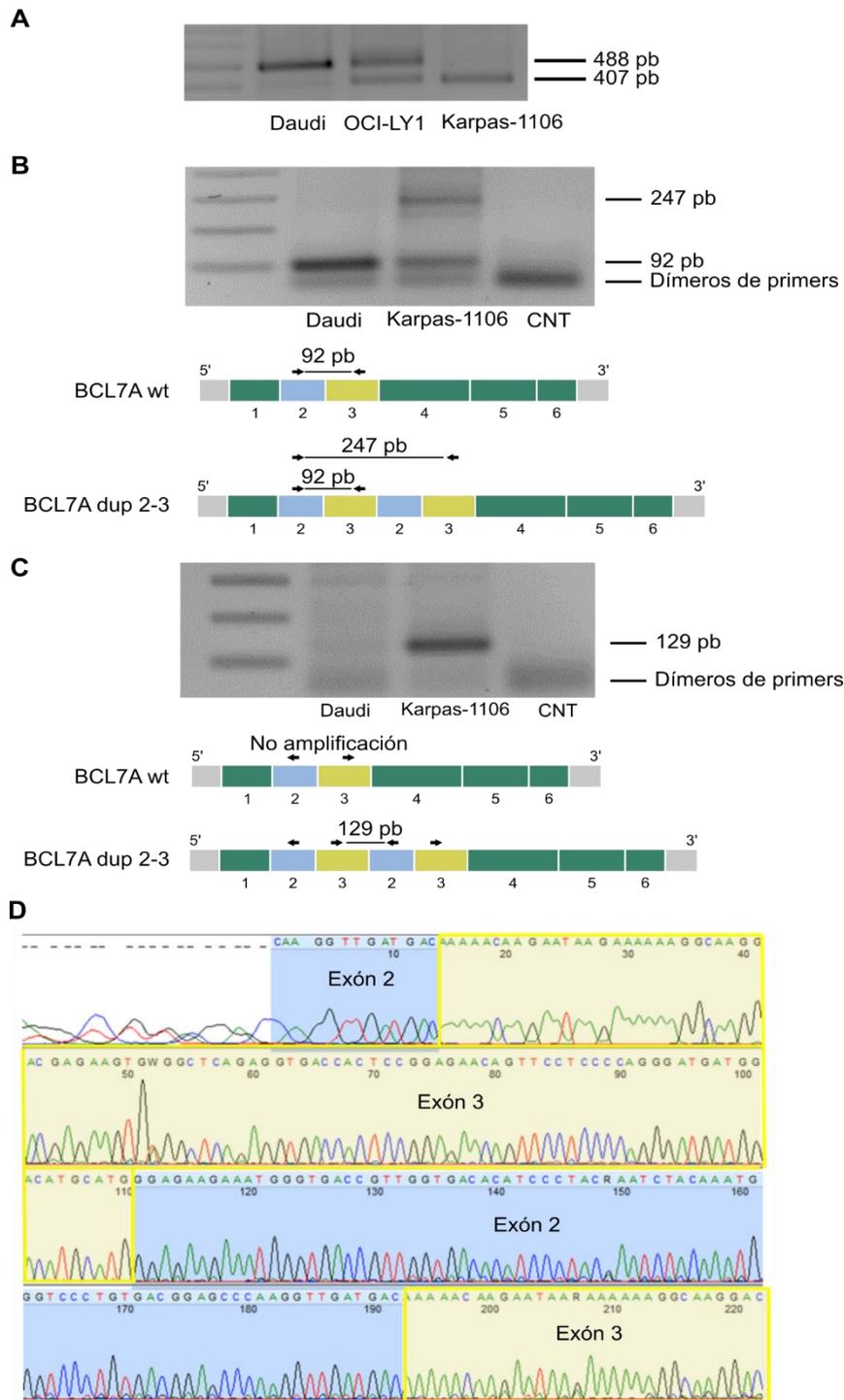


Figura 18. La línea celular Karpas-1106 presenta una duplicación en tándem que afecta a los exones 2 y 3 en un único alelo. A) Electroforesis en gel de agarosa mostrando los productos de PCR obtenidos para las líneas celulares indicadas siguiendo el esquema de PCR mostrado en la figura 17C. **B-C)** Se utilizaron diferentes diseños de PCR para la caracterización de la duplicación presente en Karpas-1106 a nivel de ARNm. La línea celular Daudi se utilizó como control para la expresión de BCL7A silvestre y se incluyó un control negativo (CNT) que no incluía molde para cada una de las PCRs. Las flechas indican la ubicación de los cebadores directos (\rightarrow) o reversos (\leftarrow). También se muestran los tamaños de banda teóricos para cada producto de PCR posible. **D)** Cromatograma obtenido mediante secuenciación de Sanger del amplicón de 247pb. Se observa que los exones 2 (resaltado en azul) y 3 (resaltado en amarillo) se encuentran duplicados en tándem.

Para la línea celular Karpas-1106, al amplificar la región entre el exón 1 y el exón 4, no se obtuvo ninguna banda de tamaño “silvestre” (**Figura 16B**). Este resultado se replicó al realizar una nueva amplificación (**Figura 18A**). Por tanto, mientras que un alelo presenta la mutación de *splicing* que genera el transcrito mutante $\Delta 27$ -BCL7A, debe haber otro tipo de variación genética que afecta al otro alelo e impide la expresión de la forma silvestre de BCL7A. Tras realizar diferentes amplificaciones en la línea celular Karpas-1106, se observó un patrón de bandas inesperado (**Figura 18B**). Al secuenciar el amplicón de tamaño aberrante (247pb) pudimos caracterizar una duplicación en tándem de los exones 2 y 3 (**Figura 18D**). Posteriormente, se confirmaron estos resultados realizando otra amplificación a nivel de ADNc. Esta amplificación podría darse exclusivamente si la duplicación mencionada tiene lugar. En este caso, se usó un cebador directo que anillaba en el exón 3 y un cebador reverso que anillaba en el exón 2 (**Figura 18C**). De esta forma, se confirmó la duplicación en tándem de los exones 2 y 3 en la línea celular Karpas-1106.

5.4 EXPRESIÓN DE BCL7A A NIVEL DE PROTEÍNA

Una vez caracterizadas las mutaciones encontradas en las distintas muestras analizadas, el siguiente paso fue determinar las consecuencias de dichas mutaciones sobre la expresión de BCL7A a nivel de proteína. Uno de los objetivos de este estudio era encontrar una línea o líneas celulares en las que se haya perdido la expresión de la proteína BCL7A silvestre, de forma que puedan usarse como modelo para restaurar la expresión de BCL7A y reevaluar el fenotipo tumoral.

Las líneas celulares candidatas después de nuestro análisis mutacional fueron OCI-LY1, Karpas-1106 y VAL. Las tres líneas celulares presentaban una mutación de *splicing* que genera un transcrito aberrante a nivel de ARNm. Además, todas ellas presentan un patrón de inactivación bialélica, donde una segunda variación genética impide la expresión de BCL7A silvestre. Basándonos en los resultados mutacionales, las líneas celulares OCI-LY1 y Karpas-1106 solo expresarían la forma mutante $\Delta 27$ -BCL7A, mientras que la línea VAL expresaría un BCL7A mutante en la posición R11 (R11T-BCL7A) junto al mutante $\Delta 27$ -BCL7A. Debido a que BCL7A expresa dos isoformas, corta y larga, podrían coexistir hasta cuatro transcritos BCL7A de distintos tamaños en una línea celular como VAL. Dos transcritos serían relativos a las variantes silvestres corta y larga y otros dos transcritos serían relativos a las variantes $\Delta 27$ -BCL7A corta y larga. En cuanto al mutante R11T-BCL7A tendría un tamaño indistinguible respecto a la forma silvestre de BCL7A (WT). En base a todo lo expuesto, en la **figura 19** se muestra un esquema de

los distintos transcritos y proteínas que, de forma teórica, podríamos encontrar en las 3 líneas celulares candidatas a usarse como modelo.

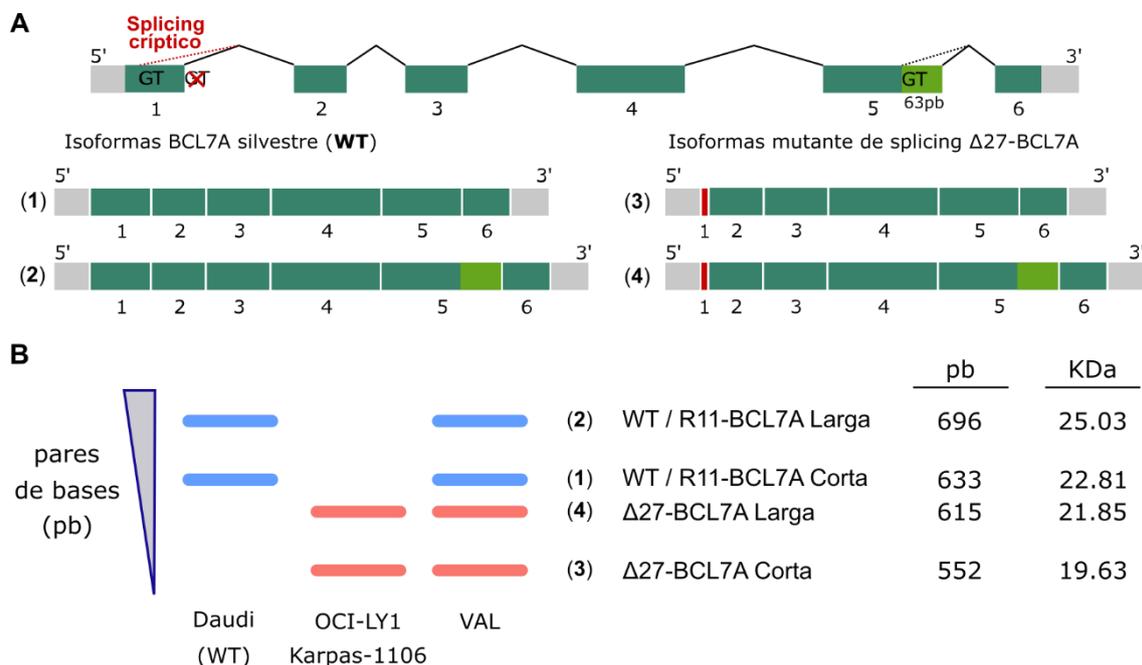


Figura 19. Representación de los posibles transcritos obtenidos a partir de BCL7A silvestre o del mutante Δ27-BCL7A. **A)** Representación de la estructura de los genes *BCL7A* y *Δ27-BCL7A* a nivel de ARNm. Los exones están coloreados en verde y las regiones UTR coloreadas en gris. La región diferencial en el exón 5 entre las isoformas cortas y largas se resalta en verde claro. La secuencia donadora de *splicing* que está mutada en la variante *Δ27-BCL7A* está tachada y el nuevo sitio de *splicing* críptico se muestra dentro del propio exón 1. Los distintos transcritos se han enumerado entre paréntesis para facilitar su visualización en el panel inferior. **B)** Representación en forma de dibujo en el que se indica cómo se ordenarían por tamaño, en pares de bases, los distintos transcritos de las variantes de *BCL7A* en las líneas celulares OCI-LY1 y Karpas-1106 en comparación a la línea celular VAL. Además, se incluye también la línea celular Daudi como referencia de línea celular que expresa *BCL7A* WT. Los valores de KDa que se muestran en la figura han sido calculados de forma teórica utilizando una herramienta bioinformática del servidor *ExpASy*: (https://web.expasy.org/compute_pi/).

Primero, se decidió analizar la expresión de *BCL7A* mediante Western blot en las líneas celulares OCI-LY1 y Karpas-1106 en comparación con las líneas celulares Daudi y SU-DHL-4, que expresan un *BCL7A* WT y se utilizaron como controles positivos. Además, usamos como controles negativos algunas de las líneas celulares para las cuales habíamos detectado una expresión nula de *BCL7A* mediante RT-qPCR (**Figura 8**). Aunque las variaciones son muy sutiles en cuanto a tamaño, podemos ver como las líneas celulares OCI-LY1 y Karpas-11060 mostraban una banda de un menor tamaño que las líneas celulares control (**Figura 20A**). Aun así, se visualiza un patrón de 3 bandas tanto para las líneas mutantes como para las líneas WT, algo que no encaja con lo esperado (**Figura 19B**). En cuanto a las líneas celulares utilizadas como controles negativos, HBL1 y KM-H2, mostraron también una banda. Estos resultados sugirieron que esa banda, de tamaño

intermedio entre las isoformas cortas y largas, podría ser una banda no específica. La única línea celular que no muestra ninguna banda en el análisis por Western blot es la línea celular NB4.

A continuación, aprovechando que NB4 mostró una expresión nula de BCL7A a nivel de proteína, se utilizó esta línea celular para expresar las diferentes variantes de BCL7A de forma exógena y verificar el patrón de bandas observado para cada una de las 4 variantes posibles entre BCL7A y $\Delta 27$ -BCL7A (**Figura 20B**). Podemos ver que las isoformas cortas y largas de las variantes WT y $\Delta 27$ -BCL7A tienen una mayor separación entre ellas de lo esperado (comparar **Figuras 19B y 20B**). Además, la expresión exógena de las variantes de BCL7A en NB4 nos confirmó que la banda intermedia es probablemente una banda no específica. Por otro lado, decidimos expresar las variantes de BCL7A también en la línea celular OCI-LY1. Si nos fijamos en la intensidad relativa de las bandas y tenemos en cuenta la propia expresión basal de las variantes $\Delta 27$ -BCL7A en OCI-LY1, el patrón de bandas es idéntico al obtenido en NB4.

Finalmente, para salir de dudas con respecto a la banda no específica, decidimos adquirir una línea celular HAP1-BCL7A-KO donde la expresión de BCL7A está inhibida mediante CRISPR-Cas9 (Ver Materiales y métodos; sección Western blot). Junto a la línea noqueada se usó como referencia la línea celular HAP1 Parental, que expresa BCL7A WT. De esta forma, se pudo validar que la banda intermedia es una banda no específica reconocida por el anticuerpo utilizado (**Figura 20C**). Teniendo en cuenta todos los Western blots y comprobaciones realizadas, en la **figura 20D** se muestra un diagrama final en el que aparecen las posiciones relativas de las distintas bandas observadas y su relación con la variante de BCL7A a la que representan.

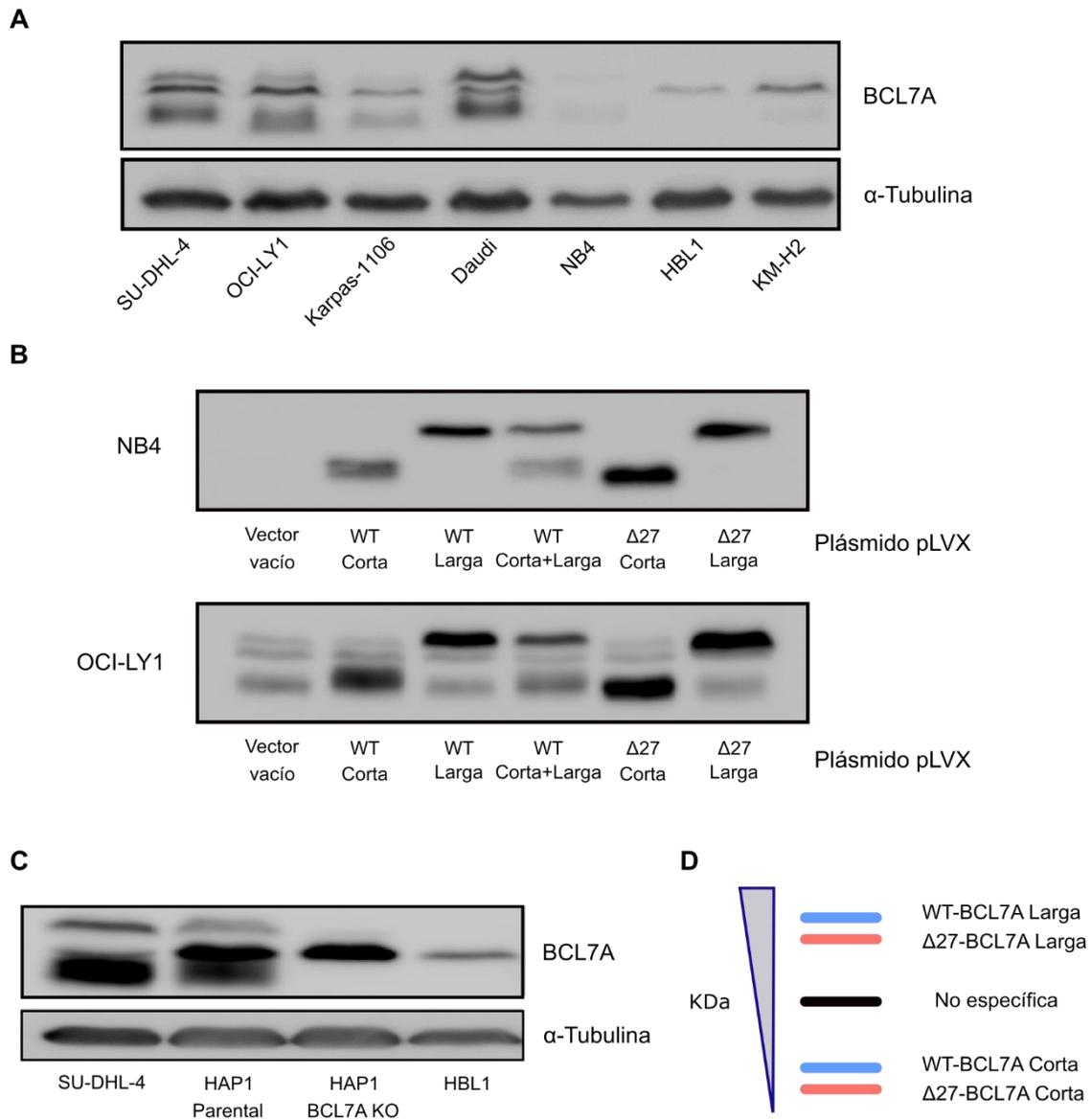


Figura 20. Expresión de BCL7A a nivel de proteína en distintos contextos celulares. A) Análisis por Western blot de la expresión endógena de BCL7A en extractos de proteína total de distintas líneas celulares. La proteína α -Tubulina se usó como control de carga. **B)** Análisis por Western blot de la expresión de BCL7A en extractos de proteína total de las líneas celulares NB4 (arriba) y OCI-LY1 (abajo) transducidas con las construcciones pLVX-IRES-ZsGreen1 que contienen las distintas isoformas corta y larga tanto de BCL7A WT como del mutante Δ 27-BCL7A o un vector vacío. **C)** Análisis por Western blot para determinar que la banda intermedia es una banda reconocida por el anticuerpo de forma no específica. Para validar el anticuerpo se utilizaron las líneas celulares HAP1 BCL7A *knockout* (KO) y HAP1 Parental. La línea celular SU-DHL-4 se utilizó como control positivo para la expresión de BCL7A silvestre. HBL1 se usó como control negativo que carece de la expresión de BCL7A. **D)** Se muestra un dibujo para representar la posición relativa de cada una de las distintas bandas observadas en los distintos Western blots realizados en este estudio y la asociación de cada una de ellas con la variante de BCL7A que representan.

Por último, una vez validada la expresión de BCL7A por Western blot, también decidimos evaluar la expresión de BCL7A en algunas líneas celulares adicionales. Lo que hicimos fue ver la expresión endógena de la proteína BCL7A en un grupo de líneas celulares derivadas de linfoma de Burkitt

y compararla con un grupo de líneas celulares derivadas de mieloma múltiple. Pues bien, se puede apreciar una clara diferencia en cuanto a la expresión de BCL7A ambos grupos de líneas celulares (**Figura 21**). Mientras que las líneas celulares derivadas de linfoma de Burkitt expresan altos niveles de BCL7A, las líneas de mieloma múltiple tienen baja o incluso nula expresión de BCL7A. Estos datos concuerdan con la idea de que la expresión de BCL7A se produce en células B que experimentan la reacción de CG, pero la expresión de BCL7A se pierde cuando las células B abandonan el CG para convertirse en una célula plasmática.

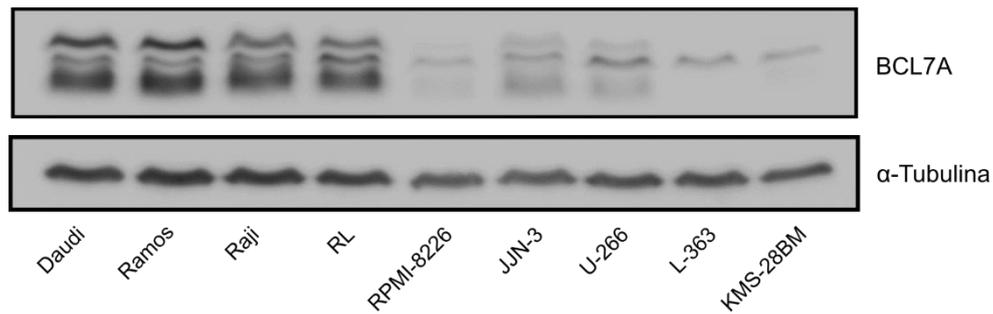


Figura 21. Expresión de BCL7A a nivel de proteína en líneas celulares derivadas de linfoma de Burkitt o mieloma múltiple. Análisis por Western blot de la expresión endógena de BCL7A en extractos de proteína total de las distintas líneas celulares que se indican. La proteína α -Tubulina se usó como control de carga.

5.5 ESTUDIO MUTACIONAL DE *BCL7A* EN COHORTES EXTERNAS

Poco después de concluir nuestro estudio mutacional, se publicaron dos artículos independientes en los que habían conseguido aglutinar un número de pacientes de DLBCL que superaba el medio millar en ambos casos y habían aplicado técnicas de secuenciación masiva para analizar exoma completo. A continuación, se detallan los datos mutacionales obtenidos para *BCL7A* en los dos estudios de DLBCL más grandes realizados hasta la fecha: el estudio de Reddy et al. con 1001 pacientes¹³⁶ y el estudio de Schmitz et al. con 574 pacientes de DLBCL¹³⁷.

Por un lado, a finales del año 2017, se publicó un artículo en la revista *Cell* que englobaba un exhaustivo análisis en 1001 pacientes de DLBCL¹³⁶. El análisis integraba datos mutacionales, datos de expresión y datos clínicos. Gracias al uso de 400 muestras pareadas tumor/línea germinal, los autores establecieron una lista de 150 genes conductores del cáncer con prevalencias mutacionales que iban desde un 1 a un 28,8%. En el caso de *BCL7A*, este gen mostraba una prevalencia mutacional del 4,5%, situándolo en la posición 65 del total de 150 genes. Sin embargo, nos dimos cuenta de que las mutaciones de *splicing* no habían sido tenidas en cuenta ya que el estudio se había restringido al análisis de mutaciones exónicas.

Por otro lado, en el año 2018 se publicó otro artículo en la revista *The New England Journal of Medicine* donde se realizó un análisis que integraba datos mutacionales, datos de expresión y

datos clínicos de 574 pacientes de DLBCL¹³⁷. En este caso, también se definía a *BCL7A* como uno de los genes frecuentemente mutados en pacientes de DLBCL, con una prevalencia mutacional de 6.1%. De nuevo, la mayoría de las mutaciones de *splicing* de *BCL7A* se habían dejado fuera del análisis debido a los criterios de filtrado aplicados por los autores del trabajo.

En el estudio mutacional que habíamos realizado nosotros sobre un total de 94 muestras que englobaban líneas celulares y pacientes de DLBCL, habíamos observado que la mutación de *splicing* era la más frecuente. Además, se caracterizó molecularmente que las mutaciones de *splicing* generaban una variante mutante de *BCL7A* que podría tener implicaciones funcionales sobre el fenotipo tumoral. Por lo tanto, quisimos obtener los datos crudos de secuenciación de los dos estudios de DLBCL recién publicados y así analizar si las mutaciones de *splicing* representan una señal de identidad mutacional en pacientes de DLBCL.

Tras obtener los permisos necesarios, se consiguió acceso a los datos crudos de WES de ambos estudios y aplicamos un método propio de análisis para determinar la tasa mutacional del gen *BCL7A*, incluyendo también las posiciones de *splicing*. El reanálisis de ambas bases de datos permitió definir 28 y 7 mutaciones de *splicing* en las bases de datos de Reddy et al. y Schmitz et al. respectivamente (**Tabla 4**). Con el fin de confirmar las nuevas mutaciones detectadas, se realizó una inspección visual de todas las muestras mutadas en *BCL7A* mediante el software *Integrative Genomics Viewer* (IGV). Al realizar este análisis, se observó un posible sobrefiltrado de mutaciones tanto en la posición del codón de inicio de la traducción como en el aminoácido 11 (R11) en *BCL7A* en el estudio de Reddy et al. (**Tabla 4**).

Tabla 4. Nuevas mutaciones identificadas para *BCL7A* en cohortes externas

Reddy et al (2017), N = 1001			
Mutación (hg38)	Localización	Nº casos	IDs de pacientes
chr12:122022092_A/G	Codón de inicio	1	2211
chr12:122022092_A/T	Codón de inicio	1	3712
chr12:122022093_T/G	Codón de inicio	2	2166, 2203
chr12:122022123_G/C	R11T	1	2478
chr12:122022124_G/T	R11S	1	2047
chr12:122022184_G/C	<i>Splicing</i>	7	2137, 2533, 2832, 3659, 3816, 3849, 3916
chr12:122022184_G/T	<i>Splicing</i>	6	2127, 2475, 2639, 2669, 2933, 3589
chr12:122022184_G/A	<i>Splicing</i>	12	2047, 2087, 2151, 2166, 2203, 2278, 2554, 2678, 3712, 3785, 3817, 3856
chr12:122022185_T/C	<i>Splicing</i>	1	3647
chr12:122022185_T/G	<i>Splicing</i>	2	2076, 2707
Schmitz et al (2018), N = 574			
Mutación (hg38)	Localización	Nº casos	IDs de pacientes
chr12:122022184_G/C	<i>Splicing</i>	4	DLBCL11655, DLBCL10945, DLBCL11558, DLBCL11242
chr12:122022184_G/T	<i>Splicing</i>	2	DLBCL10863, DLBCL10514
chr12:122022184_G/A	<i>Splicing</i>	1	DLBCL11584

Las localizaciones cromosómicas de las mutaciones vienen indicadas en base a la versión del genoma GRCh38/hg38. Los identificadores IDs de los pacientes son los que se han utilizado en cada uno de los estudios. Las mutaciones en el codón de inicio y en la posición R11 se rescataron tras inspeccionar visualmente en IGV los archivos de secuenciación de los pacientes en los que detectamos mutaciones de *splicing* u otras mutaciones inactivadoras previamente descritas en *BCL7A* por los autores originales de los estudios.

Nuestro análisis tuvo una limitación para poder determinar el estado somático de las nuevas mutaciones detectadas, ya que nosotros no tuvimos acceso a los datos de secuenciación de las muestras pareadas de tejido sano de las que disponían los autores originales de los estudios. Para determinar si las nuevas mutaciones definidas por nuestro análisis eran somáticas, primero observamos que las mutaciones en el codón de inicio o en R11 en *BCL7A* no se encuentran en cohortes de más de 100.000 sujetos sin cáncer en la base de datos de *gnomAD* (<https://gnomad.broadinstitute.org>). Asimismo, el resto de mutaciones en el codón de inicio y

en R11 que se definen en el estudio de Reddy et al. están confirmadas como somáticas. En cuanto a las mutaciones de *splicing*, se observó que las mutaciones en el primer sitio de *splicing* de *BCL7A* están presentes en la base de datos *gnomAD* a frecuencias alélicas por debajo de $1/10^5$ ($N = 145.172$ alelos secuenciados en la posición c.92+1). Además, encontramos evidencia externa de mutaciones somáticas en la posición c.92+1 de *BCL7A* en tres muestras de linfoma de células B en la base de datos COSMIC y una muestra en la base de datos de TCGA-DLBC (**Tabla 5**). En cuanto a las dos mutaciones en la posición c.92+2 que sí estaban presentes en los datos originales del estudio de Schmitz et al. (2018), se confirmaron como somáticas en el estudio original. En conjunto, estos hechos parecen indicar que, a pesar de la falta de información de la línea germinal en la mayoría de los pacientes analizados, es altamente probable que las mutaciones recién identificadas en *BCL7A* sean somáticas.

Tabla 5. Mutaciones de *BCL7A* en la posición c92+1 definidas como somáticas en bases de datos

Base de datos	Muestra ID	Mutación ID	Tipo de muestra
COSMIC	GHE1438	COSV55850485	DLBCL
COSMIC	N10	COSV55850485	Linfoma de zona marginal
COSMIC	QC2-40-T2	COSV55846678	DLBCL
TCGA	TCGA-FF-8062	-	DLBCL

COSMIC (v91) y TCGA-DLBC (N = 22 muestras pareadas tumor/normal)

5.6 LAS MUTACIONES EN *BCL7A* SE CONCENTRAN EN EL EXÓN 1 Y PRESENTAN UN PATRÓN DE INACTIVACIÓN BIALÉLICA

Al combinar ambas cohortes externas de pacientes ($N=1575$ pacientes totales), la inclusión de las mutaciones identificadas en nuestro análisis aumentó la frecuencia de mutación de *BCL7A* en un 30%, pasando de un 5,1% (80/1575) a un 6,6% (104/1575) de pacientes con mutaciones en *BCL7A*. Estos números representan únicamente las mutaciones en secuencias codificantes o sitios de *splicing*. Si nos fijamos en la localización de las mutaciones con respecto a la secuencia codificante del gen, la mayoría de las mutaciones de *BCL7A* se encuentran en el primer exón o en el sitio donador de *splicing* del intrón 1 (**Figura 22**).

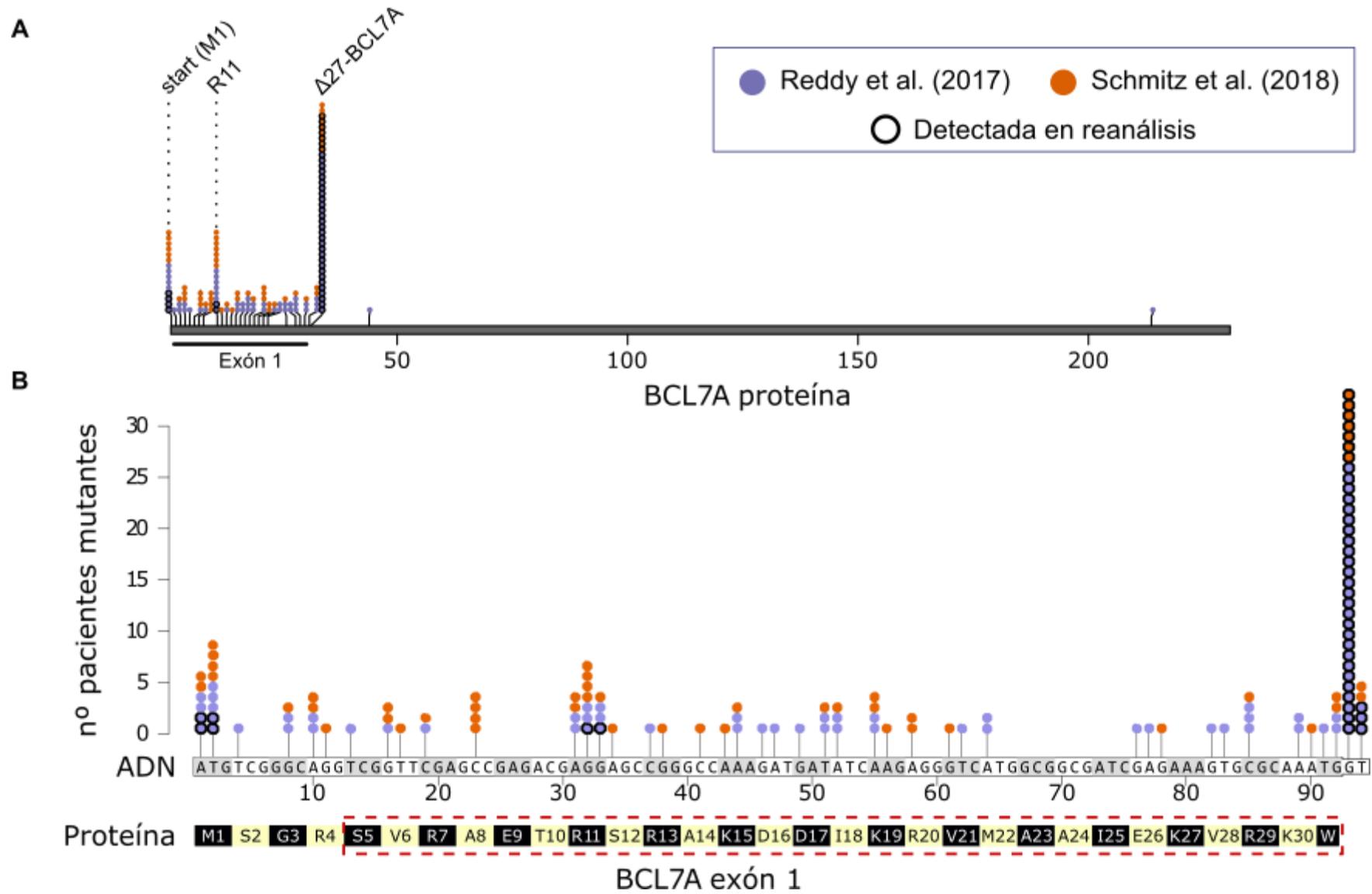


Figura 22. Gráfico de mutaciones en *BCL7A* en cohortes externas de DLBCL. Se muestran las mutaciones combinadas de las cohortes de Reddy et al. (2017) y Schmitz et al. (2018). Los círculos delineados en negro representan las nuevas mutaciones encontradas en nuestro reanálisis. **A)** Gráfico *lollipop* de mutaciones en *BCL7A* a nivel de proteína a lo largo de toda la secuencia de la proteína *BCL7A*. Cada círculo representa una muestra mutada. **B)** Gráfico *lollipop* de mutaciones en *BCL7A* a nivel de ADN en el exón 1 y el primer sitio de donador de *splicing*. Los aminoácidos delecionados en la variante mutante $\Delta 27$ -*BCL7A* (S5_W31del) se muestran delimitados por un cuadro rojo de línea discontinua y representan el dominio que hemos nombrado como BCL7_Nt.

De entre todas las mutaciones, la mutación de *splicing* es un punto caliente mutacional (del inglés “*Hot-spot mutation*”) ya que acumula un gran número de mutaciones. Además, se realizó un test estadístico para evaluar otros posibles puntos calientes en la secuencia codificante de *BCL7A*. Para ello, se utilizó un modelo estadístico que asume como hipótesis nula una probabilidad de mutación uniforme en todos los aminoácidos del primer exón de *BCL7A*, incluyendo el sitio de *splicing* como si fuese un aminoácido más. A continuación, se realizó una prueba binomial para evaluar si la frecuencia observada de mutaciones en cada aminoácido era mayor que la esperada bajo la hipótesis nula. Pues bien, además del punto caliente mutacional ya mencionado en el sitio donador de *splicing* ($p=8,24 \cdot 10^{-25}$), se caracterizaron otros dos puntos calientes, aunque menos significativos, que concentran un alto porcentaje de mutaciones. Se trata del codón R11 y del codón de inicio (ambas posiciones con $p=1,20 \cdot 10^{-4}$) (**Figura 23A**).

En secciones posteriores mostraremos porqué las mutaciones de *splicing* son consideradas como mutaciones inactivadoras o de pérdida de función. Por el momento, y para simplificar la nomenclatura, a partir de ahora nos referiremos conjuntamente a mutaciones en el sitio de *splicing*, mutaciones con cambio de marco de lectura, mutaciones de pérdida del codón de inicio de la traducción y mutaciones sin sentido en *BCL7A* como “mutaciones inactivadoras”. De tal modo, las mutaciones inactivadoras junto con las mutaciones con cambio de sentido en el codón R11 engloban más de la mitad (73/135) de todas las mutaciones en *BCL7A* (**Figura 23A**). A continuación, al analizar la coocurrencia de mutaciones en *BCL7A*, se observó que un total de 15 pacientes presentaban más de una mutación inactivadora y/o mutación en R11 (**Tabla 6**). En todos estos casos, excepto en dos, estas mutaciones estaban presentes en diferentes alelos (**Figura 23B**). En cambio, no se observaron patrones de exclusión alélica con mutaciones con cambio de sentido distintas a R11 o M22L, aunque son solo dos pacientes presentaban la mutación M22L.

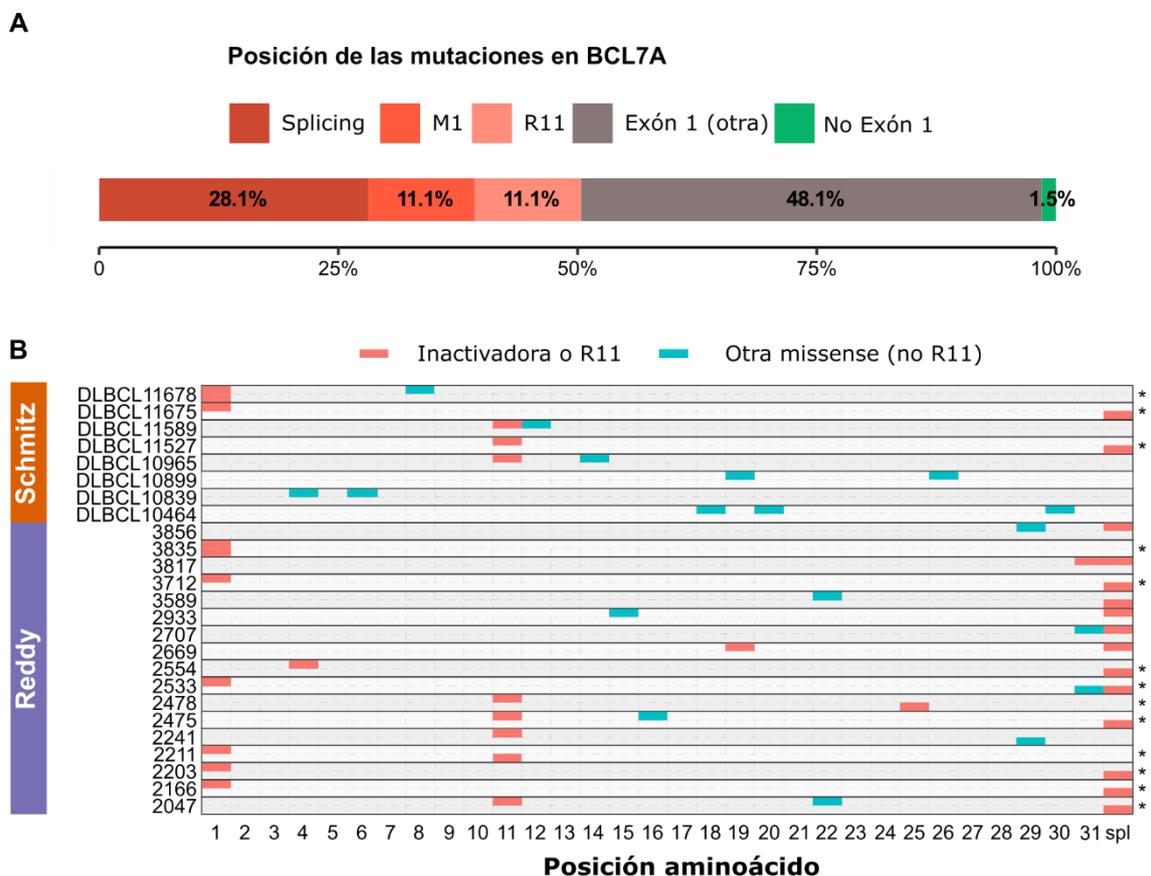


Figura 23. Patrón mutacional de BCL7A en cohortes externas de DLBCL e inactivación bialélica.

A) Distribución de mutaciones en el gen *BCL7A* después de combinar las cohortes de los estudios de Reddy et al. (2017) y Schmitz et al. (2018). Los tres puntos calientes mutacionales significativos según nuestro análisis se muestran en distintas tonalidades de rojo. **B)** Patrón de inactivación bialélica en *BCL7A* en muestras de pacientes con DLBCL. Los identificadores de paciente (IDs) de las cohortes de Schmitz y Reddy se muestran en el eje Y. Para cada paciente, cada uno de los dos alelos se representa como una fila individual y los diferentes aminoácidos + sitio donador de *splicing* (spl) se representan en el eje X. Los asteriscos en el borde derecho indican aquellos pacientes que muestran un patrón de inactivación bialélica, donde hay coexistencia de mutaciones inactivadoras y/o mutaciones en R11 (coloreadas en rojo) en diferentes alelos. El resto de mutaciones con cambio de sentido diferentes a las mutaciones en R11 están coloreadas en azul.

Tabla 6. Análisis de inactivación bialélica en pacientes de DLBCL que tienen más de una mutación en *BCL7A*
Reddy et al (2017), N = 1001

IDs Pacientes	Mutaciones en <i>BCL7A</i>	¿Están en el mismo alelo?
2047	R11S + M22L + <i>splicing</i>	No (<i>splicing</i> vs otras) Sí (R11S y M22L)
2166	Codón de inicio + <i>splicing</i>	No
2203	Codón de inicio + <i>splicing</i>	No
2211	Codón de inicio + R11G	No
2241	R11S + R29C	No
2475	R11M + D16N + <i>splicing</i>	No (<i>splicing</i> vs otras) Sí (R11M y D16N)
2478	R11T + I25fs	No
2533	Codón de inicio + K30R + <i>splicing</i>	No (Codón de inicio vs otras) Sí (<i>splicing</i> y K30R)
2554	R4fs + <i>splicing</i>	No
2669	K19X + <i>splicing</i>	Sí
2707	W31G + <i>splicing</i>	Sí
2933	K15R + <i>splicing</i>	Sí
3589	M22L + <i>splicing</i>	No
3712	Codón de inicio + <i>splicing</i>	No
3817	W31X + <i>splicing</i> (doble SNV)	Sí
3835	Codón de inicio + Codón de inicio	No
3856	R29C + <i>splicing</i>	Sí

Continúa en la siguiente página

Schmitz et al (2018), N = 574		
IDs Pacientes	Mutaciones en <i>BCL7A</i>	¿Están en el mismo alelo?
DLBCL11675	Codón de inicio + <i>splicing</i>	No
DLBCL11589	R11S + S12G (doble SNV)	Sí
DLBCL11527	R11G + deleción del sitio donador de <i>splicing</i>	No
DLBCL10899	K19E + E26D	Sí
DLBCL10839	R4G + V6L	Sí
DLBCL11678	Codón de inicio + Codón de inicio + A8G	No (Ambas mutaciones del codón de inicio) Sí (uno de los codones de inicio y A8G)
DLBCL10965	R11T + A14V	Sí
DLBCL10464	I18V + R20G + K30N	Sí

Los identificadores IDs de los pacientes son los que se han utilizado en cada uno de los estudios. Los IDs en color azul corresponden a los 15 pacientes en los que coexisten más de una mutación inactivadora y/o mutación en R11. Las mutaciones en el codón de inicio y en R11 marcadas en rojo son aquellas que fueron rescatadas mediante análisis visual en IGV (Ver **tabla 4**). “doble SNV”: Los SNVs afectan a nucleótidos consecutivos. “fs”: *frameshift* (mutación con cambio en el marco de lectura).

Por último, se realizó un análisis para poder inferir si las mutaciones en *BCL7A* han podido generarse en etapas tempranas o tardías del desarrollo tumoral. Para ello, se analizaron las frecuencias de los alelos mutantes para *BCL7A*. Una mayor frecuencia alélica se asocia a la aparición de mutaciones en etapas tempranas. De tal forma, se observó que la frecuencia alélica para los alelos mutantes era $\geq 30\%$ en más de la mitad (43/73) de las mutaciones inactivadoras o mutaciones en R11, e incluso eran $\geq 70\%$ para 3 de las mutaciones de *splicing* (resultados no mostrados). Además, estos valores se han obtenido sin tener en cuenta la pureza o heterogeneidad tumoral. En definitiva, el patrón de inactivación bialélica sumado a la alta frecuencia mutacional de las mutaciones inactivadoras apoyan la idea de que *BCL7A* pueda ejercer un papel supresor tumoral en DLBCL.

5.7 LA ACCIÓN DE LA ENZIMA AID NO JUSTIFICA LA TOTALIDAD DE MUTACIONES QUE PRESENTA *BCL7A*.

Estudios anteriores han mostrado que *BCL7A* es diana de la enzima deaminasa de citidina inducida por activación (*AID* o *Aicda*)¹⁴³ y que las mutaciones en *BCL7A* presentan la huella característica de la acción de AID en linfomas derivados del CG^{144,154}. En las células B normales del CG, la enzima AID genera una alta tasa de mutaciones en genes de inmunoglobulina a través del mecanismo de SHM²⁷. Sin embargo, se puede producir SHM de forma aberrante en locus de genes distintos a los de inmunoglobulinas, generando mutaciones que contribuyen a la inestabilidad genómica y linfomagénesis en células B¹⁵⁵. En este trabajo, hemos aplicado secuenciación de nueva generación (NGS) para analizar mutaciones inducidas por AID en *Bcl7a* en células B de CG de placas de Peyer de ratón¹⁵⁶ y así poder evaluar la implicación directa de AID en el patrón mutacional de *Bcl7a*. Brevemente, se comparó la frecuencia de mutaciones transicionales en pares de bases de citosina/guanina (C/G) en células B de CG de ratones noqueados para la enzima *Aicda*, y de ratones *Ung*^{-/-}*Msh2*^{-/-}. Las enzimas UNG y MSH2 se respaldan mutuamente para reparar las lesiones inducidas por AID¹⁴⁸, pero en ausencia de UNG y MSH2, las deaminaciones inducidas por la AID no se procesan y dan lugar a transiciones C→T y G→A, la huella representativa de la acción de AID^{157,158}.

En primer lugar, se utilizó ADN genómico de células B de CG para amplificar una región de 790pb y el producto de PCR fue analizado por NGS. El posterior análisis de secuenciación de *Bcl7a* se limitó a una región de 312pb que estaba cubierta por más de 100 lecturas de alta calidad. La región de 312pb incluye la región 5'-UTR y el primer exón de *Bcl7a* (**Figura 24A**). Posteriormente, analizamos la frecuencia de mutaciones transicionales en C/G dentro de las secuencias WRCY/RGYW, que son los puntos calientes mutacionales más representativos de la acción de AID¹⁵⁹. Estas secuencias se mostraron altamente mutadas en *Bcl7a* en células B que expresan AID en comparación con las células B deficientes para AID (**Figura 24B**). Por otro lado, la frecuencia de mutaciones transicionales era mucho mayor en C/G dentro de los dominios (W)RCY o (R)GYW en comparación a la observada en C/G fuera de los dominios mencionados.

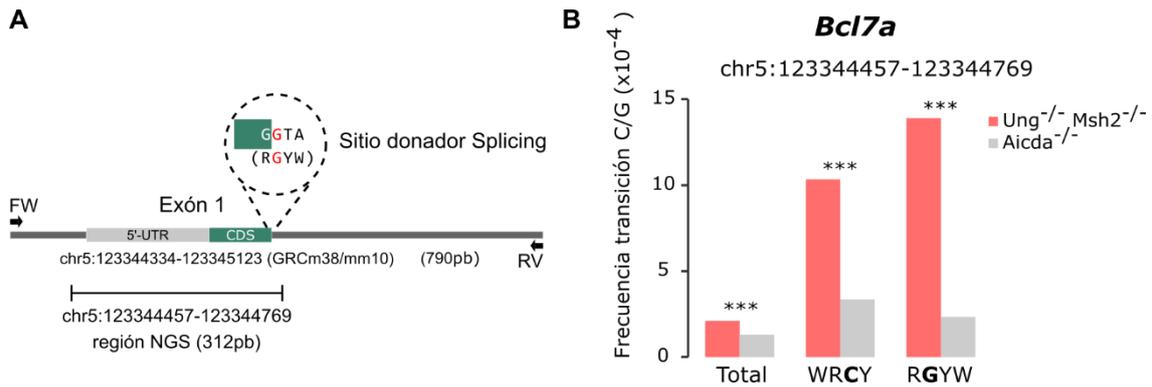


Figura 24. Mutaciones en *Bcl7a* debido a la acción de AID. A) Representación gráfica de la región genómica amplificada mediante PCR y posteriormente secuenciada por NGS. Esta región abarca la 5'-UTR y el primer exón de *Bcl7a*. Se indican las coordenadas genómicas del amplicón completo (790pb; entre el cebador directo (Fw) y reverso (Rv)) y de la región analizada después de realizar el filtrado de calidad (312pb, ver sección Materiales y métodos). **B)** Frecuencia de mutaciones transicionales en los nucleótidos C/G totales y en C/G dentro de las secuencias reconocidas por AID: WRC(Y)/(R)GYW (W = A/T; R = G/A; Y = C/T); en células B de CG de las placas de Peyer de ratones *Ung*^{-/-}*Msh2*^{-/-} y *Aicda*^{-/-}. Prueba de Fisher de dos colas; ***, p < 10⁻⁴.

Para validar la metodología empleada y a modo de control, se analizó de la frecuencia de mutaciones transicionales en la región de 312pb en dos motivos diferentes que no son reconocidos por AID: las secuencias CTCT y GGCG; donde las letras subrayadas denotan las citosinas susceptibles de sufrir mutaciones transicionales. La diferencia entre ratones *Ung*^{-/-}*Msh2*^{-/-} y *Aicda*^{-/-} no fue significativa para ninguno de los dos motivos analizados (**Figura 25** y **Tabla 7**).

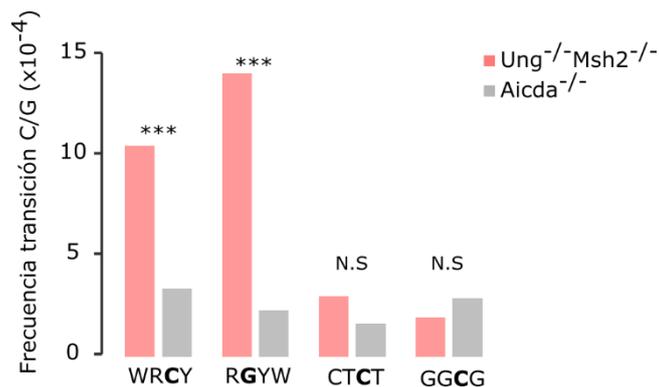


Figura 25. Mutaciones inducidas por AID en secuencias no específicas. Frecuencia de mutaciones transicionales en nucleótidos C/G en secuencias WRC(Y)/(R)GYW en comparación con C/G en dos secuencias no reconocidas por AID (W = A/T; R = G/A; Y = C/T). Prueba de Fisher de dos colas; ***, p < 10⁻⁴; N.S = No significativo.

Tabla 7. Cálculo de frecuencia de mutaciones transicionales en distintas secuencias

	Bases mutadas (transiciones)		Bases no mutadas		Frecuencia de mutación		
	<i>Ung/Msh2</i> <i>dKO</i>	<i>Aicda</i> <i>KO</i>	<i>Ung/Msh2</i> <i>dKO</i>	<i>Aicda</i> <i>KO</i>	<i>Ung/Msh2 dKO</i>	<i>Aicda KO</i>	p-valor
WRCY	36	34	34821	101477	1.03E-03	3.35E-04	3.23E-06
RGYW	46	20	33083	85846	1.39E-03	2.33E-04	2.00E-12
CTCT	13	35	43276	214315	3.00E-04	1.63E-04	0.078
GGCG	3	11	15537	37890	1.93E-04	2.90E-04	0.767

Se muestran los datos calculados de frecuencia transicional en C/G en las secuencias mostradas. Se muestran los valores individuales para cada uno de los modelos de ratón. Prueba de Fisher de dos colas.

A continuación, se evaluó la correlación entre las mutaciones encontradas en *Bcl7a* en el modelo de ratón y las mutaciones identificadas en *BCL7A* en pacientes con DLBCL. A pesar de que estamos comparando datos de NGS obtenidos en ratón con mutaciones en muestras humanas, las secuencias entre el codón de inicio y el sitio de *splicing* son idénticas al 98% entre especies. Los nucleótidos con la frecuencia neta de transición más alta en ratones AID⁺ presentan, generalmente, mutaciones en muestras humanas de DLBCL (**Figura 26**, panel superior). Sin embargo, una alta frecuencia de mutaciones transicionales en nuestro modelo de ratón no se correlacionaba positivamente con una alta prevalencia de mutaciones en *BCL7A* en muestras humanas de DLBCL (**Figura 26**, panel inferior). Además, las secuencias del codón de inicio de la traducción o el codón R11 no coinciden con secuencias de reconocimiento por AID y tienen una frecuencia de transición nula en nuestro modelo de ratón. Por otro lado, el motivo RGYW en el primer sitio de *splicing* está altamente mutado en pacientes con DLBCL, pero tiene una frecuencia de transiciones relativamente baja en nuestro modelo de ratón. En conjunto, estas observaciones sugieren que la acción de AID no explica completamente el patrón mutacional encontrado en *BCL7A* en los tumores primarios de DLBCL. Nuestros datos también apoyan que las mutaciones en el sitio *splicing* proporcionan una ventaja de crecimiento en las células tumorales, ya que están sobrerrepresentadas en comparación con el amplio espectro de mutaciones inducidas por AID que se observan en las células CG B no transformadas.

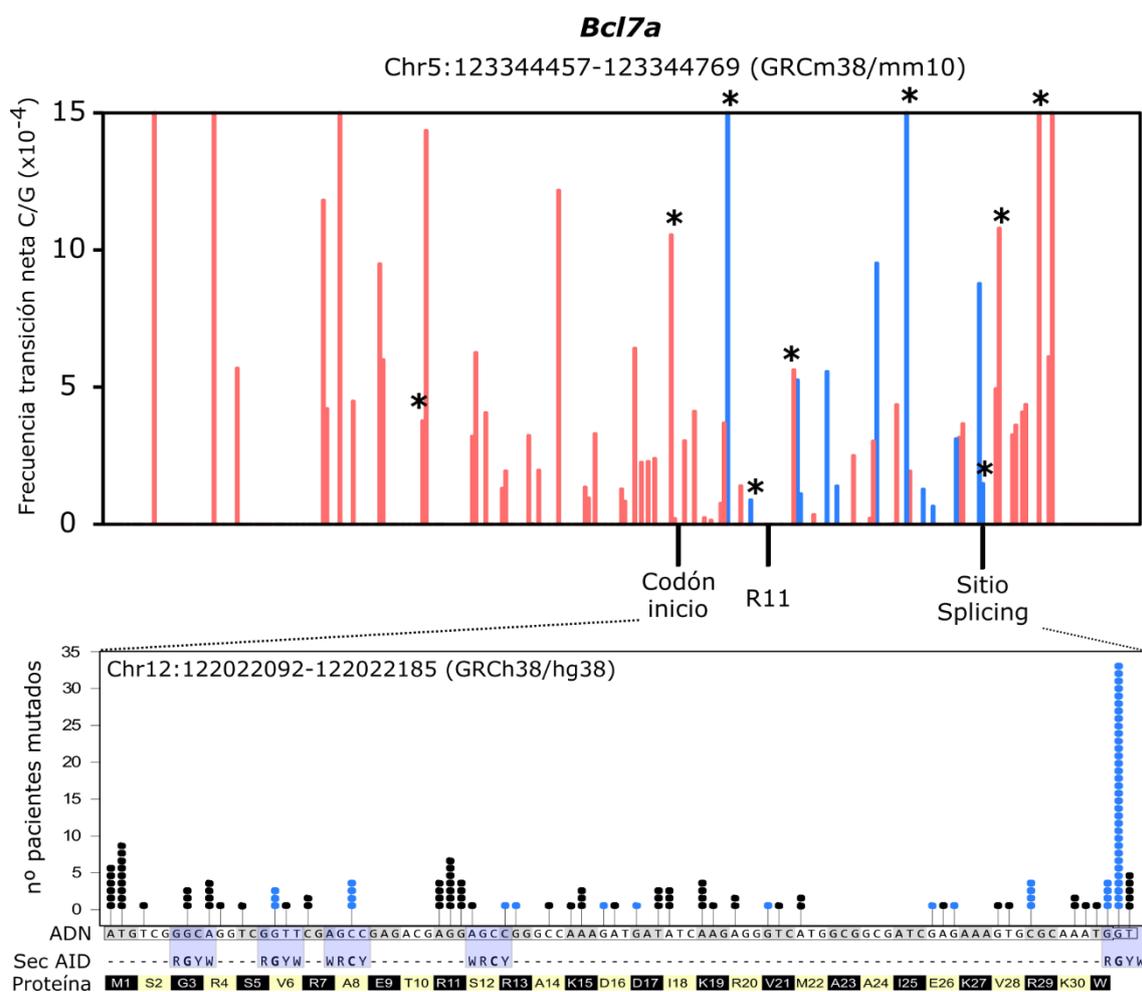


Figura 26. Correlación entre las mutaciones inducidas por AID en nuestros modelos de ratón y mutaciones en pacientes con DLBCL. Panel superior) Análisis base por base en el amplicón de 312pb cuyas coordenadas genómicas aparecen indicadas. El gráfico muestra la frecuencia de transición neta (TS_freq) de mutaciones en C/G (TS_freq neta = TS_freq Ung/Msh2 dKO - TS_freq Aicda KO). Las posiciones de nucleótidos que se correlacionan con las mutaciones encontradas en las muestras de DLBCL se han coloreado en azul y los asteriscos se han colocado en nucleótidos C/G que pertenecen a los motivos de reconocimiento por AID WRCY/RGYW. En el borde inferior del panel superior se indican las posiciones relativas a los puntos calientes mutacionales (codón de inicio, R11 y sitio donador de *splicing*) en muestras humanas de DLBCL. **Panel inferior)** Gráfico “lollipop” de mutaciones en *BCL7A* a nivel de ADN en la región que se extiende desde el codón de inicio hasta el primer sitio donador de *splicing*. Las mutaciones que se muestran son las encontradas de forma conjunta en las cohortes de Reddy et al. (2017) y Schmitz et al. (2018). Los círculos azules indican aquellas mutaciones que coinciden con los nucleótidos representados con picos azules en el panel superior. Se muestran las secuencias de reconocimiento de AID (Sec AID) WRCY/RGYW, resaltadas en azul, así como la secuencia de *BCL7A* a nivel de proteína.

5.8 EL DOMINIO AMINO-TERMINAL DE *BCL7A* ES CRÍTICO PARA SU INTERACCIÓN CON EL COMPLEJO SWI/SNF.

El exón 1 de *BCL7A* alberga la mayoría de las mutaciones. Es precisamente el exón 1 de *BCL7A* el que se deletiona casi totalmente debido a las mutaciones de *splicing* caracterizadas en este estudio. A esa región deletionada, la hemos denominado “dominio BCL7_Nt”. Este dominio está

conservado evolutivamente y se comparte con los otros dos miembros de la familia de genes *BCL7*: *BCL7B* y *BCL7C*¹²⁷. Teniendo en cuenta que los tres genes *BCL7* son subunidades excluyentes de los complejos SWI/SNF, el dominio amino-terminal podría tener una función común en integrar las subunidades *BCL7* al complejo SWI/SNF. De tal forma, nosotros postulamos que las mutaciones de *splicing* impiden la unión de *BCL7A* al complejo SWI/SNF mediante la eliminación del dominio *BCL7_Nt*. Para evaluar esta hipótesis, se realizaron estudios de inmunoprecipitación de *BCL7A* en células HBL1. Este modelo celular carece de la expresión de *BCL7A* (**Figura 20**) y fue utilizado para expresar de forma exógena las distintas variantes de *BCL7A*. Así, las células HBL1 fueron transducidas tanto con *BCL7A* silvestre como con el mutante $\Delta 27$ -*BCL7A*. A continuación, se analizó si *SMARCA4*, la subunidad catalítica del complejo SWI/SNF, inmunoprecipitaba junto a *BCL7A*. Se observó que mientras que *BCL7A* silvestre se unía a *SMARCA4*, el mutante $\Delta 27$ -*BCL7A* no lo hacía (**Figura 27A**).

Posteriormente, con el objetivo de describir más a fondo la composición del complejo SWI/SNF residual y determinar qué subunidades podrían permanecer unidas tanto a *BCL7A* silvestre como al mutante $\Delta 27$ -*BCL7A*, llevamos a cabo un análisis de cromatografía líquida acoplada a espectrometría de masas (LC-MS/MS). Prácticamente todas las subunidades del complejo SWI/SNF se detectaron en los inmunocomplejos que contenían *BCL7A* silvestre. En cambio, ninguna de las subunidades del complejo SWI/SNF permanecieron unidas a la forma mutante $\Delta 27$ -*BCL7A* (**Figura 27B**). A continuación, nos preguntamos si la exclusión de *BCL7A* del complejo SWI/SNF modificaba las interacciones entre *SMARCA4* y las subunidades centrales del complejo SWI/SNF. Mediante nuevos análisis de inmunoprecipitación, pudimos comprobar que el estado mutacional de *BCL7A* no afectaba a la unión de *SMARCA4* a las subunidades centrales del complejo SWI/SNF, ya que estas co-inmunoprecipitaron con *SMARCA4* en células que expresaban tanto *BCL7A* silvestre como el mutante $\Delta 27$ -*BCL7A* (**Figura 27C**). En conjunto, estos resultados muestran que el dominio *BCL7_Nt*, que se pierde debido a mutaciones en el sitio de *splicing*, es necesario para que *BCL7A* se una al complejo SWI/SNF, pero la ausencia de *BCL7A* parece no afectar a la integridad del complejo residual (**Figura 27D**).

En los análisis mostrados hasta ahora, solamente se han tenido en cuenta las subunidades del propio complejo SWI/SNF que se unen a las subunidades *BCL7A* o *SMARCA4*. En cambio, a partir de los datos de LC/MS-MS podemos plantearnos la siguiente pregunta: ¿qué otras proteínas han inmunoprecipitado con las variantes *BCL7A* salvaje o con el mutante $\Delta 27$ -*BCL7A*? En el **apéndice IV** se muestra una tabla de proteínas que han sido detectadas mediante espectrometría de

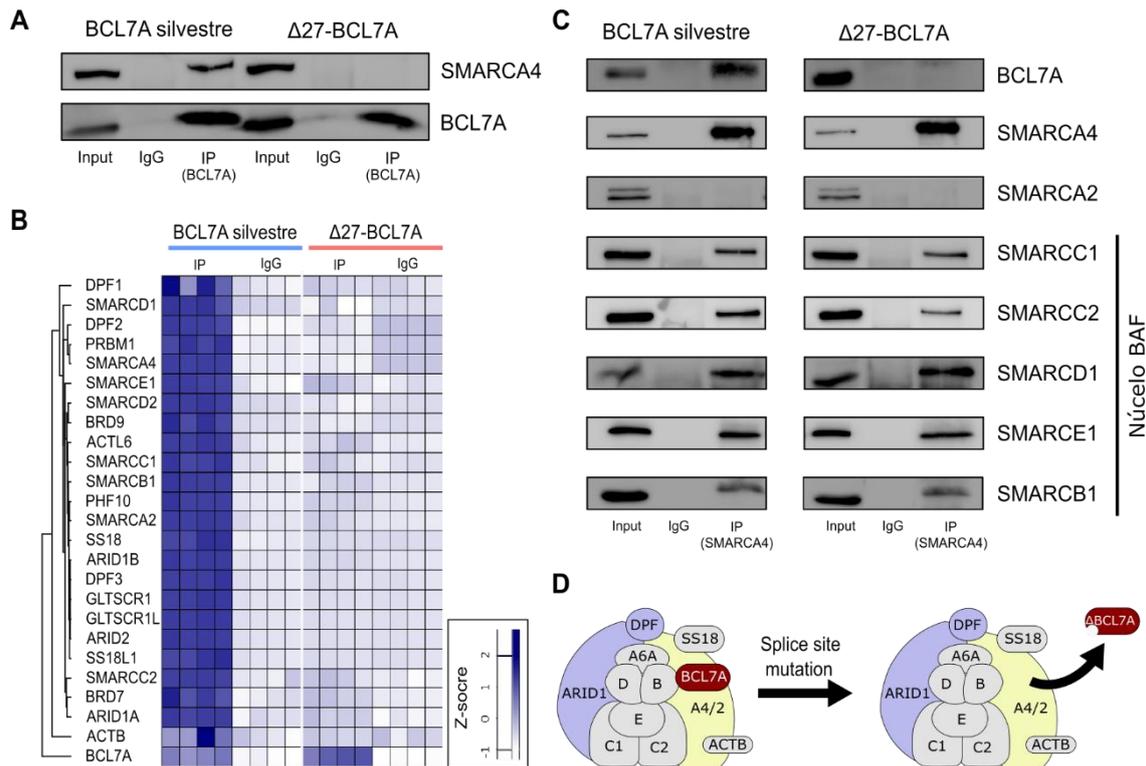


Figura 27. La pérdida del dominio BCL7A_Nt impide la unión de BCL7A al complejo SWI/SNF.

A) Inmunoprecipitación usando un anti-BCL7A a partir de extractos de proteína total de células HBL1 transducidas con BCL7A silvestre o $\Delta 27$ -BCL7A (isoformas cortas). La co-inmunoprecipitación de las subunidades BCL7A y SMARCA4 se evaluó mediante Western blot. El gel mostrado es una imagen representativa después de realizar el experimento dos veces. **B)** Análisis por LC/MS-MS de las subunidades del complejo SWI/SNF. Los inmunocomplejos se purificaron usando un anti-BCL7A en células HBL1 transducidas con BCL7A silvestre o $\Delta 27$ -BCL7A (isoformas cortas). Se realizaron dos réplicas biológicas para cada condición (IP e IgG) y cada réplica se corrió dos veces. Los valores de “Label-free Quantification” (LFQ) se representaron después de la normalización por Z-Score. Todas las subunidades representadas en el mapa de calor se caracterizaron como interactores de BCL7A al mostrar los siguientes valores; $p < 0.05$ y $\log_2 FC > 3$. Las subunidades DPF1 y ACTB fueron las dos únicas subunidades con un $\log_2 FC < 3$ (1.4 y 1.8 respectivamente). **C)** Los inmunocomplejos analizados en este apartado se purificaron usando un anti-SMARCA4 a partir de extractos de proteína total de células HBL1 transducidas con BCL7A silvestre o $\Delta 27$ -BCL7A (isoformas cortas). Se analizó por Western blot la co-inmunoprecipitación de las subunidades centrales del complejo SWI/SNF, así como de BCL7A, SMARCA4 y SMARCA2. **D)** Representación del complejo residual BAF. La pérdida del dominio BCL7A_Nt, debido a las mutaciones de *splicing*, implica que la proteína resultante $\Delta 27$ -BCL7A no puede unirse al complejo SWI/SNF. Las subunidades del complejo BAF se abreviaron de la siguiente forma: Las letras A4/2, B, C1, C2, D y E representan las diferentes subunidades 'SMARC'; ARID1 = ARID1A/B; DPF = DPF1/2/3; SS18 = SS18 o SS18L1.

masas en las diferentes condiciones analizadas. Se han considerado como interactores todas aquellas proteínas que muestran una proporción entre la muestra y su control IgG con valores \log_2 superiores a 1,5 ya sea en células HBL1 transducidas con la forma salvaje de BCL7A o la forma mutante $\Delta 27$ -BCL7A. De forma similar a las proteínas del complejo SWI/SNF, la mayoría de las proteínas detectadas pierden la capacidad de unirse a la forma mutante $\Delta 27$ -BCL7A en

comparación con la forma salvaje de BCL7A. Probablemente, las uniones a esas proteínas no se producen a partir del propio BCL7A, sino que son proteínas que se unen al resto del complejo SWI/SNF. En cambio, hay algunas proteínas que parecen unirse de forma similar a las diferentes variantes de BCL7A (**Apéndice IV**). En principio podríamos asumir que esas uniones se producen a través de secuencias de BCL7A distintas al dominio BCL7_Nt. Aun así, hay que tener cuidado con estas apreciaciones ya que también podría tratarse de proteínas que se unen de forma inespecífica al anticuerpo usado para inmunoprecipitar. De hecho, casi con total seguridad, la proteína PAWR se corresponde a la banda no específica que aparece en los Western blot llevados a cabo en este estudio (**Figura 20**). Esto lo sabemos ya que, hemos llevado a cabo análisis de inmunoprecipitación en la línea celular NB4 (resultados no mostrados) y la proteína PAWR no se detectó mediante espectrometría de masas siguiendo la misma metodología que para la línea celular HBL1. NB4 era la única línea celular para la que no se detectaba la banda no específica en los Western blots.

Por otro lado, se realizó un análisis mediante la herramienta STRING (<https://string-db.org/>) para analizar si las múltiples proteínas detectadas mediante LC/MS-MS muestran enriquecimiento en determinadas vías funcionales. De tal forma, algunos de los componentes celulares más representados fueron miembros del complejo remodelador de la cromatina Mi-2/NuRD o miembros del complejo relacionado con la actina Arp2/3 (**Apéndice IV**). Todas estas proteínas apenas se detectaban al inmunoprecipitar el mutante $\Delta 27$ -BCL7A. Esto no quiere decir que su unión dependa del dominio BCL7_Nt, sino que, probablemente, reflejan proteínas que se asocian al complejo SWI/SNF en el contexto celular analizado.

5.9 LA RESTAURACIÓN DE LA EXPRESIÓN DE BCL7A TIENE UN EFECTO SUPRESOR TUMORAL

Hasta el momento hemos identificado la alta prevalencia de las mutaciones de *splicing* de BCL7A en muestras de DLBCL. Además, hemos caracterizado cómo esta mutación impide la unión de la proteína resultante $\Delta 27$ -BCL7A al complejo SWI/SNF. Para poder evaluar el efecto de restaurar la expresión de BCL7A, decidimos usar como modelo aquellas líneas celulares derivadas de DLBCL que no expresan la proteína BCL7A silvestre, sino que expresan la forma mutante $\Delta 27$ -BCL7A. Teniendo en cuenta que la capacidad de proliferar de manera descontrolada es una de las características que define a las células tumorales, lo primero que quisimos evaluar es el cambio de proliferación en nuestros modelos celulares tras restaurar la expresión de BCL7A. Con este fin, desarrollamos un ensayo de crecimiento celular competitivo en células OCI-LY1, que

expresan exclusivamente el mutante $\Delta 27$ -BCL7A, y células VAL, que co-expresan el mutante $\Delta 27$ -BCL7A y el mutante R11T-BCL7A.

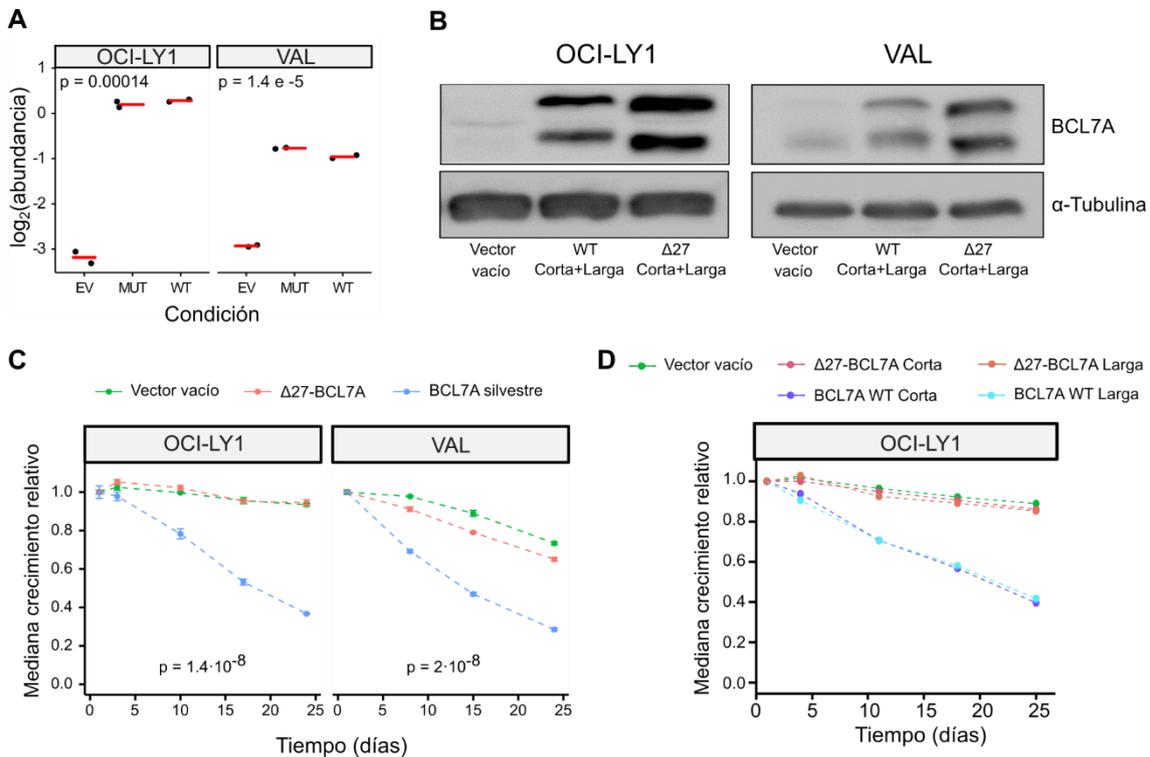


Figura 28. La restauración de la expresión de BCL7A induce un fenotipo supresor tumoral *in vitro*.

A) Análisis de la expresión de BCL7A mediante RT-qPCR en las líneas celulares OCI-LY1 y VAL transducidas con un vector vacío (EV), $\Delta 27$ -BCL7A (MUT) o BCL7A silvestre (WT). Se analizaron dos réplicas biológicas de transducciones independientes y cada una por triplicado. La abundancia relativa de los transcritos, convertida a logaritmo, se calculó utilizando el método de curva estándar y usando la medida de expresión de GAPDH para normalizar. La línea roja horizontal representa la media. Se muestran los valores p de ANOVA. **B)** Análisis de BCL7A mediante Western blot utilizando extractos de proteína de las líneas celulares OCI-LY1 (panel izquierdo) y VAL (panel derecho) transducidas con las construcciones SAR2(LTR)-IRES-ZsGreen1 que expresan las diferentes isoformas de BCL7A silvestre, $\Delta 27$ -BCL7A o el vector vacío. El análisis de α -Tubulina se utilizó como control de carga. **C)** Ensayo de crecimiento celular competitivo en las líneas celulares de DLBCL transducidas con (i) vector vacío, (ii) $\Delta 27$ -BCL7A o (iii) BCL7A silvestre. Tanto para la forma silvestre como para el mutante $\Delta 27$ -BCL7A se co-expresaron las isoformas corta y larga. Se muestran los valores p de ANOVA para las comparaciones a tiempo final del experimento. **D)** Ensayo de crecimiento celular competitivo en la línea celular OCI-LY1 transducida con (i) vector vacío, (ii) BCL7A WT isoforma corta, (iii) BCL7A WT isoforma larga (iv) $\Delta 27$ -BCL7A isoforma corta, o (v) $\Delta 27$ -BCL7A isoforma larga. **C-D)** Las células ZsGreen1⁺ se purificaron mediante FACS 5 días después de la transducción y se cultivaron hasta obtener aproximadamente 5×10^6 células totales. Posteriormente, las células ZsGreen1⁺ se mezclaron con células ZsGreen1⁻ para obtener tres cultivos independientes por condición con 1×10^6 células totales por cultivo y un porcentaje inicial de células ZsGreen1⁺ del 60%. La media del porcentaje de células ZsGreen1⁺ en cada punto de tiempo se normalizó con respecto a la media del porcentaje de células ZsGreen1⁺ en el día 1. Los puntos de partida (día 1) para el análisis de crecimiento celular competitivo fueron 8 y 15 días después de la transducción para las líneas celulares OCI-LY1 y VAL respectivamente.

Ambas líneas celulares fueron transducidas con diferentes construcciones lentivirales que co-expresan de forma estable las variantes de BCL7A junto con el fluorocromo ZsGreen1. Con el fin de validar nuestros modelos, la expresión de las variantes de BCL7A se analizó mediante RT-qPCR y Western blot (**Figura 28A-B**). La restauración de la expresión de BCL7A silvestre dio lugar a una disminución en la proliferación en las líneas celulares OCI-LY1 y VAL en comparación con las células transducidas tanto con el vector vacío como con el mutante $\Delta 27$ -BCL7A (**Figura 28C**). Por otro lado, la expresión independiente de las isoformas corta o larga de BCL7A silvestre o $\Delta 27$ -BCL7A también indujeron un fenotipo anti-proliferativo (**Figura 28D**), de forma equivalente al inducido por la co-expresión de ambas isoformas (**Figura 28C**). Cabe destacar que el fenotipo inducido por la expresión del mutante $\Delta 27$ -BCL7A no era significativamente diferente al fenotipo inducido por el vector vacío (prueba t de *Student* a tiempo final, $p > 0.05$). Estos datos sugieren que BCL7A desempeña un papel supresor tumoral en nuestros modelos de DLBCL, fenotipo que se pierde debido a la mutación de *splicing*.

Tras comprobar cómo la restauración de la expresión de BCL7A disminuye la capacidad tumoral de la línea celular OCI-LY1 *in vitro*, decidimos utilizar esta línea celular para realizar un xenotrasplante en ratones NSG inmunodeficientes. Para ello, se creó una línea celular estable expresando luciferasa (OCI-LY1/Luc⁺) mediante transducción lentiviral. La luciferasa es capaz de metabolizar el sustrato luciferina y generar bioluminiscencia. Las células OCI-LY1/Luc⁺ fueron transducidas con las distintas variantes de BCL7A o con el vector vacío. Posteriormente, los ratones se inyectaron intravenosamente con 5×10^6 células en tres cohortes diferentes de 10 ratones. El crecimiento tumoral se siguió *in vivo* gracias al sistema empleado. Al inyectar a los ratones intraperitonealmente con luciferina a distintos tiempos, se pudo monitorizar la bioluminiscencia, la cual es proporcional al crecimiento de las células tumorales inyectadas. Como se puede apreciar en la **figura 29**, el crecimiento tumoral fue significativamente menor para los ratones inyectados con células OCI-LY1/Luc⁺ que expresaban BCL7A silvestre en comparación con los ratones que fueron inyectados con células transducidas tanto con el vector vacío como con la forma mutante $\Delta 27$ -BCL7A. Estos resultados, junto con los datos de los ensayos de crecimiento celular competitivo *in vitro*, indican que BCL7A tiene un papel supresor tumoral en DLBCL, el cual se pierde debido a mutaciones en el sitio donador de *splicing*.

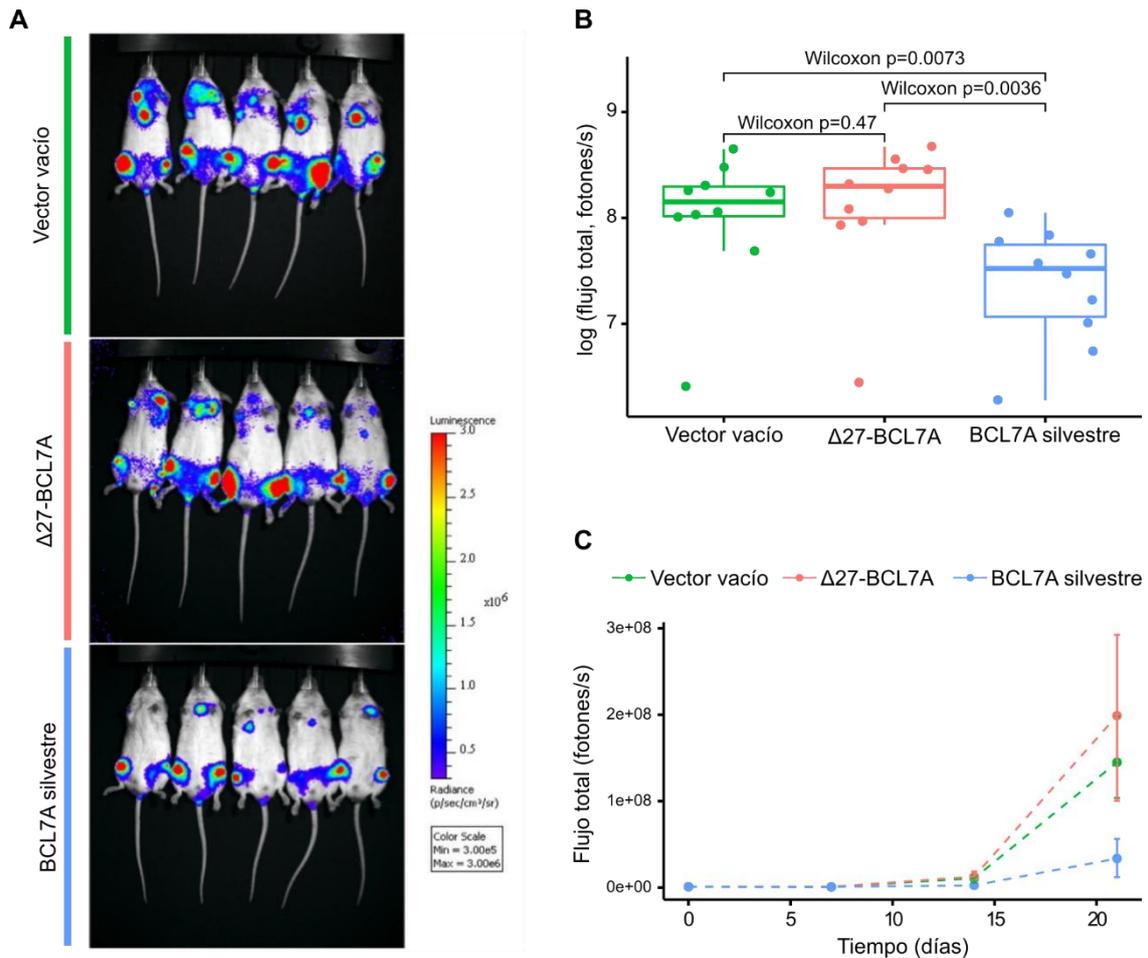


Figura 29. Xenotrasplante de la línea celular OCI-LY1 transducida con las variantes de BCL7A en ratones NSG. Se inyectaron ratones NSG con la línea celular OCI-LY1/Luc⁺ previamente transducida con las construcciones pLVX-IRES-ZsGreen1 conteniendo (i) vector vacío, (ii) $\Delta 27$ -BCL7A o (iii) BCL7A silvestre. Las células inyectadas a día 0 (12 días después de la transducción) eran sólo células positivas para el reportero ZsGreen1⁺. Se utilizaron diez ratones NSG por grupo. Se tomaron imágenes de los ratones a días 0, 7, 14 y 21 post-inyección usando el sistema de imágenes *IVIS-Spectrum*. **A)** Imágenes en posición ventral a día 21 de los 5 ratones/grupo con la señal de bioluminiscencia más alta. **B)** Señal de bioluminiscencia, transformada a logaritmo, en el día 21 después de inyectar las células (10 ratones/grupo en posición ventral). Las comparaciones por pares se realizaron utilizando una prueba de rangos con signo de Wilcoxon. **C)** Cuantificación de la bioluminiscencia (flujo total en fotones/s) en los distintos grupos ratones a lo largo del tiempo (10 ratones/grupo en posición ventral). Las barras de error representan el rango intercuartil.

5.10 BCL7A PARTICIPA EN VÍAS DE ACTIVACIÓN DE CÉLULAS B Y EN LA REGULACIÓN DE GENES IMPLICADOS EN EL CICLO CELULAR

Con el objetivo de indagar en las consecuencias funcionales derivadas de la restauración de la expresión de BCL7A, se realizó un análisis de RNA-Seq para las líneas celulares OCI-LY1 y VAL usando tres condiciones diferentes: sobreexpresión de BCL7A silvestre, sobreexpresión del mutante $\Delta 27$ -BCL7A y transducción con el vector vacío. Se obtuvieron dos réplicas biológicas para cada condición utilizada. El propio análisis de expresión diferencial sirvió como validación

del modelo experimental utilizado, al confirmar que BCL7A era el gen más sobreexpresado al comparar las condiciones BCL7A silvestre vs. vector vacío (cambio de expresión (FC) de 3,71 para OCI-LY1 y 1,61 para VAL), así como en la comparación de la expresión de $\Delta 27$ -BCL7A vs. vector vacío (FC de 3,76 para OCI-LY1 y 1,70 para VAL). En cambio, los niveles de BCL7A no variaron al comparar la sobreexpresión de BCL7A silvestre vs $\Delta 27$ -BCL7A (FC de 0,99 para OCI-LY1 y 0,95 para VAL) (Ver figura 28A).

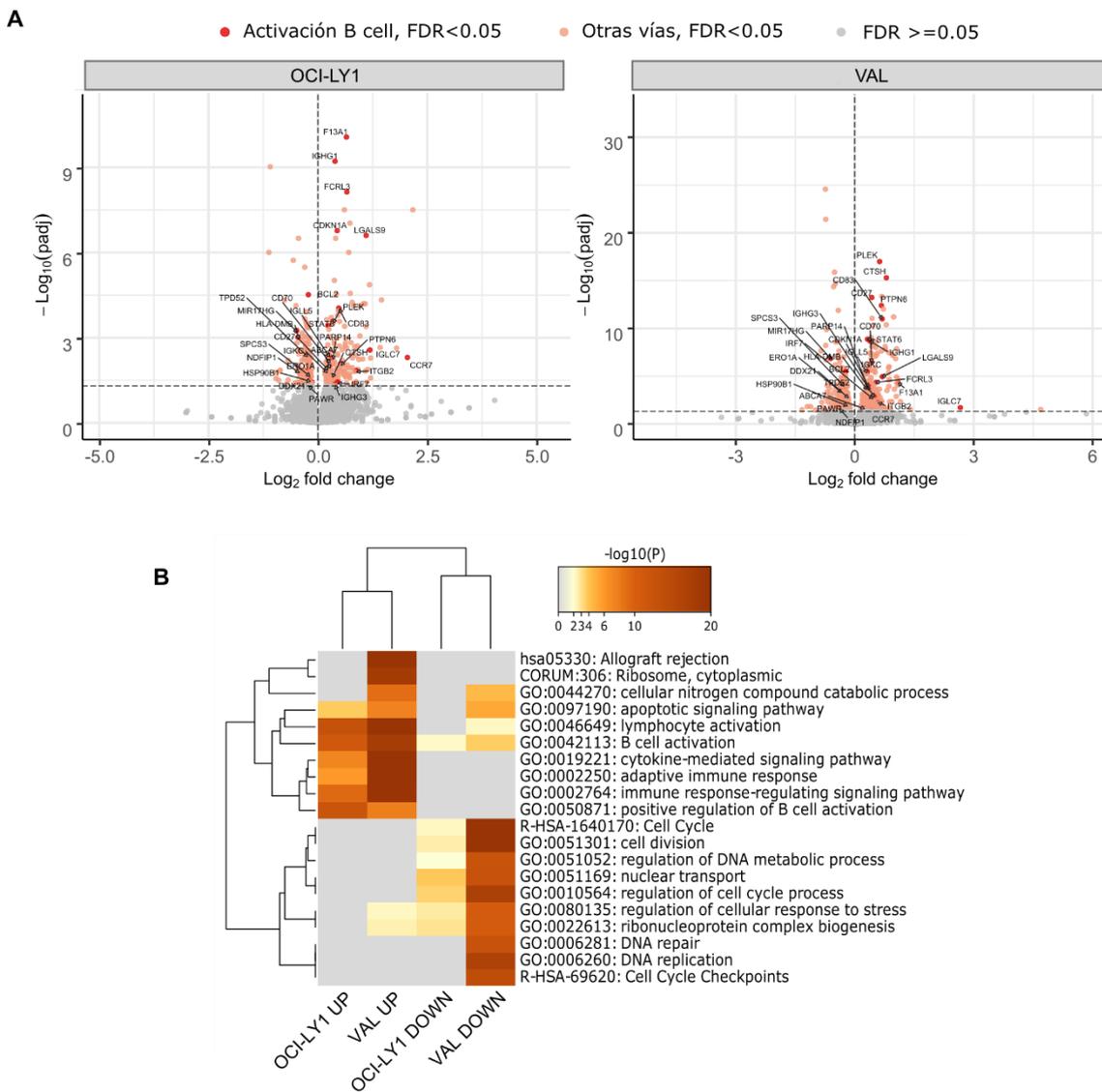


Figura 30. Análisis de expresión diferencial mediante RNA-Seq en las líneas celulares OCI-LY1 y VAL después de la sobreexpresión de BCL7A silvestre vs. sobreexpresión de $\Delta 27$ -BCL7A. A) Volcano plot destacando los genes expresados diferencialmente en las líneas celulares OCI-LY1 (izquierda) y VAL (derecha). Se muestran los nombres de los genes pertenecientes al término GO “*regulation of B cell activation*” según nuestro análisis de enriquecimiento llevado a cabo en Metascape. B) Mapa de calor que muestra los términos obtenidos al realizar un análisis de enriquecimiento de términos por ontología de genes mediante Metascape. Como input se utilizaron los genes diferencialmente expresados al comparar la sobreexpresión de BCL7A silvestre vs. sobreexpresión de $\Delta 27$ -BCL7A en ambas líneas celulares OCI-LY1 y VAL, separados en función de si estaban inducidos (UP) o inhibidos (DOWN).

Los ensayos de crecimiento celular competitivo mostraron que, al restaurar la expresión de BCL7A silvestre, la tasa de proliferación de las células tumorales disminuía. En cambio, este fenotipo no se producía al expresar la forma mutante $\Delta 27$ -BCL7A. Por tanto, decidimos comparar los modelos de sobreexpresión de BCL7A silvestre vs. $\Delta 27$ -BCL7A e identificamos 320 genes expresados diferencialmente en OCI-LY1 y 978 en VAL (valor p-ajustado (padj) < 0.05) (**Figura 30A**). Para determinar si estos genes están implicados en determinadas vías funcionales, llevamos a cabo un análisis de enriquecimiento de términos por ontología de genes, usando la opción de mapa de calor de la herramienta *Metascape*¹⁶⁰. Para ello, dividimos los genes diferencialmente expresados en genes que aparecían inducidos (*UP*) o inhibidos (*DOWN*) en ambas líneas celulares. Como podemos ver en la **figura 30B**, los genes inducidos tras restaurar la expresión de BCL7A están relacionados con vías de activación de células B. Por el contrario, los genes inhibidos muestran enriquecimiento en vías de señalización relacionadas con el ciclo celular, lo cual es consistente con la disminución en la proliferación observada tras restaurar la expresión de BCL7A tanto en los modelos *in vitro* como *in vivo* utilizados en este estudio.

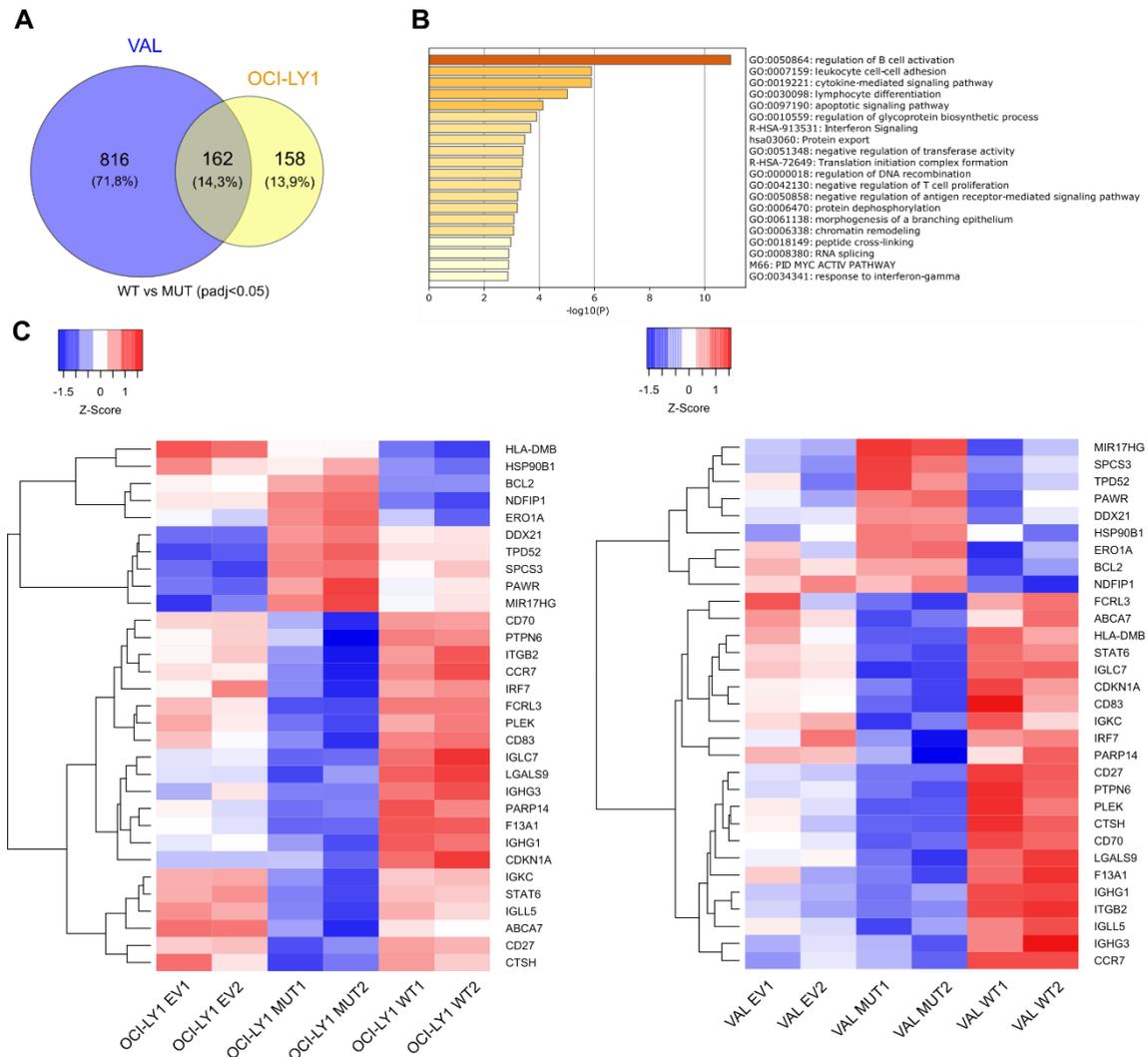


Figura 31. Los genes diferencialmente expresados de forma común en ambos modelos celulares OCI-LY1 y VAL están relacionados con vías de activación de células B. **A)** Representación, mediante diagrama de Venn, del número de genes diferencialmente expresados (padj < 0.05) en las líneas celulares OCI-LY1 y VAL al comparar la sobreexpresión de BCL7A silvestre (WT) vs. sobreexpresión del mutante $\Delta 27$ -BCL7A (MUT). 162 genes eran comunes entre ambas líneas celulares. **B)** Gráfica de barras que muestra los términos enriquecidos por ontología génica obtenidos a partir de los 162 genes diferencialmente expresados de forma común en las líneas OCI-LY1 y VAL al comparar la condición WT vs MUT. **C)** Mapas de calor que muestran los valores de expresión de 31 genes diferencialmente expresados al comparar WT vs MUT y que son comunes en los dos modelos celulares OCI-LY1 y VAL. Los 31 genes pertenecen al término GO “*regulation of B cell activation*” de acuerdo al análisis de enriquecimiento llevado a cabo en Metascape. También se muestran los valores de expresión de los 31 genes en el caso de transducir con el vector vacío (EV). Los niveles de expresión se han normalizado mediante Z-Score.

Tal y como hemos visto, al comparar los modelos de expresión BCL7A silvestre vs. $\Delta 27$ -BCL7A en ambas líneas celulares, existe cierto grado de homología en cuanto a los genes diferencialmente expresados. Más de la mitad de los genes alterados en OCI-LY1, eran comunes en ambas líneas celulares (**Figura 31A**). Para determinar si estos genes comunes se agrupan dentro de alguna vía funcional, realizamos un análisis de enriquecimiento de términos por ontología de genes, utilizando la opción de diagrama de barras de la herramienta *Metascape*. Las vías funcionales

más enriquecidas estaban relacionadas con la biología de las células B (**Figura 31B**) e incluyen: regulación de la activación de células B, adhesión inter-leucocitaria, vías de señalización mediadas por citoquinas y diferenciación de linfocitos. Entre las diferentes vías que aparecen en el análisis, la vía "regulación de la activación de células B" (GO: 00508064) destacaba con la mayor significancia ($p=1,14 \cdot 10^{-11}$) e incluye 31 genes que se expresaban diferencialmente en las líneas celulares OCI-LY1 y VAL (**Figura 31C**).

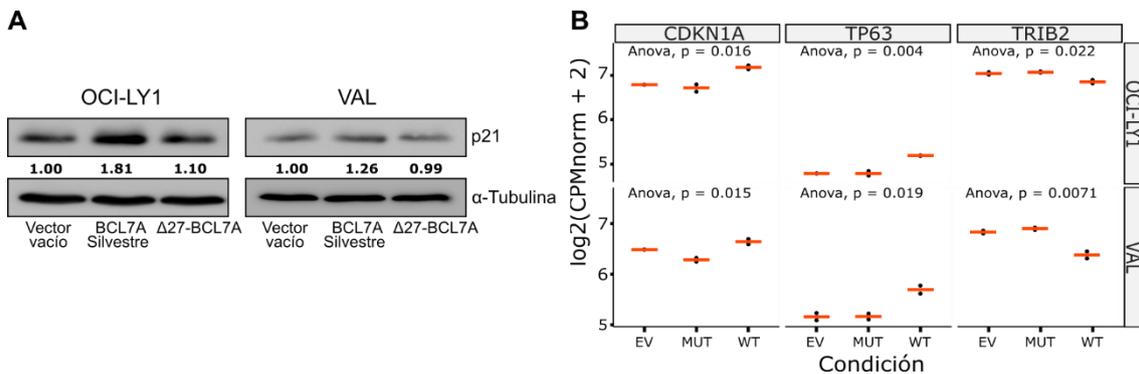


Figura 32. Expresión diferencial de determinados genes al restaurar la expresión de BCL7A. (A) Análisis mediante Western blot de p21 en extractos de proteína total de las líneas celulares OCI-LY1 (izquierda) y VAL (derecha) transducidas con las construcciones SAR2(LTR)-IRES-ZsGreen1 expresando las isoformas corta y larga tanto de BCL7A silvestre como del mutante $\Delta 27$ -BCL7A o un vector vacío. La proteína α -Tubulina se usó como control de carga y la cuantificación de bandas se llevó a cabo mediante el software ImageJ v1.49c, usando la opción "Plot Lanes". **B)** Valores de expresión de los genes *CDKN1A* (p21), *TP63* y *TRIB2* en función a nuestros datos de RNA-Seq. Las lecturas de expresión génica obtenidas y normalizadas mediante el software DESeq2 se transformaron a cuentas por millón (CPMnorm) y se transformaron a logaritmo en R escribiendo el siguiente código: `cpm(..., log = TRUE, prior.count = 2)`. Las líneas rojas horizontales representan la media. Se muestran los valores p de ANOVA.

Entre los diferentes genes que aparecen diferencialmente expresados destaca el gen *CDKN1A* (p21), cuya expresión está regulada durante la diferenciación de células B. La expresión de p21 permite un estado de proliferación más permisivo, necesario para la expansión clonal de células B en el CG^{33,161}. El análisis mediante Western blot mostró que la expresión de p21 incrementa al restaurar la expresión de BCL7A (**Figura 32A**). Otros genes relacionados con la proliferación que se expresaron diferencialmente al restaurar BCL7A fueron *TP63* o *TRIB2* (**Figura 32B**). Por un lado, el gen supresor tumoral *TP63*, el cual se ha visto que modula la transcripción de *CDKN1A*¹⁶², se mostró inducido en células que expresan BCL7A silvestre. Por otro lado, encontramos una inhibición en la expresión de *TRIB2*, que actúa como oncogén en neoplasias hematológicas¹⁶³ y que recientemente se ha caracterizado como un oncogén capaz de promover la tumorigénesis mediante el bloqueo de la vía de señalización AP4/*CDKN1A*¹⁶⁴. En conjunto, nuestros resultados sugieren que BCL7A regula la expresión de genes clave para las células B y que pueden estar vinculados con el fenotipo supresor tumoral que se observa tanto *in vitro* como *in vivo*.

5.11 EL COMPLEJO SWI/SNF ESTÁ FRECUENTEMENTE MUTADO EN EL SUBTIPO GCB-DLBCL

A continuación, quisimos evaluar si, además de *BCL7A*, otros miembros del complejo SWI/SNF presentan variaciones génicas en DLBCL. También, se estudió la posible coocurrencia de mutaciones entre *BCL7A* y otros genes relevantes en el contexto del DLBCL. Para ello, reanalizamos las cohortes utilizadas por Reddy et al. (2017)¹³⁶ y Schmitz et al. (2018)¹³⁷. Primero, nos centramos exclusivamente en el complejo SWI/SNF y consideramos todas las subunidades descritas para los distintos complejos SWI/SNF en humanos: BAF, PBAF y ncBAF; con un total de 31 subunidades^{75,116} (**Apéndice I**), e incluimos las mutaciones de *BCL7A* detectadas en nuestro propio análisis. De esta forma, se definieron un total de 289/1001 (29%) pacientes mutantes para el complejo SWI/SNF en la cohorte de Reddy et al. y 238/574 (41,5%) en la cohorte de Schmitz et al. Todos ellos mostraron mutaciones en al menos un gen del complejo SWI/SNF (**Figura 33** y **Tabla 8**). *BCL7A*, junto a los genes *ARID1A*, *ARID1B*, *SMARCA4* y *ACTB*, fueron los genes del complejo SWI/SNF mutados con mayor frecuencia en pacientes de DLBCL. De hecho, estos 5 genes fueron los únicos que Reddy et al. definieron como mutados en su cohorte. Posteriormente, se realizó un análisis de coocurrencia de mutaciones entre las distintas subunidades del propio complejo SWI/SNF. Se encontró que no existía ninguna relación significativa de coocurrencia o exclusión mutua de mutaciones entre cualquier par de subunidades del complejo SWI/SNF (prueba exacta de Fisher, $p > 0.05$; **Figura 33**).

Por otro lado, se analizó la distribución de las mutaciones de los distintos genes del complejo SWI/SNF entre los distintos subtipos de DLBCL. En primer lugar, las mutaciones en *BCL7A* afectaban preferentemente a las muestras de GCB-DLBCL en comparación con ABC-DLBC. Dependiendo de la cohorte analizada los porcentajes de mutación de *BCL7A* fueron de 8,8-13,4% en pacientes GCB, 3,8-4,4% en pacientes ABC, y 5,3-6,1% en pacientes no clasificados. Aunque esta tendencia se observó para varios genes SWI/SNF, sólo fue estadísticamente significativa para *BCL7A* (**Tabla 8**). Ahora bien, al considerar el complejo SWI/SNF como un todo y clasificar las muestras como mutantes o no mutantes en función de si tienen al menos una mutación en alguno de los genes del complejo SWI/SNF, observamos que las mutaciones en el complejo SWI/SNF se producen preferentemente en el subtipo GCB-DLBCL (prueba exacta de Fisher, $p < 0.05$; para ambas cohortes analizadas; **Tabla 8**).

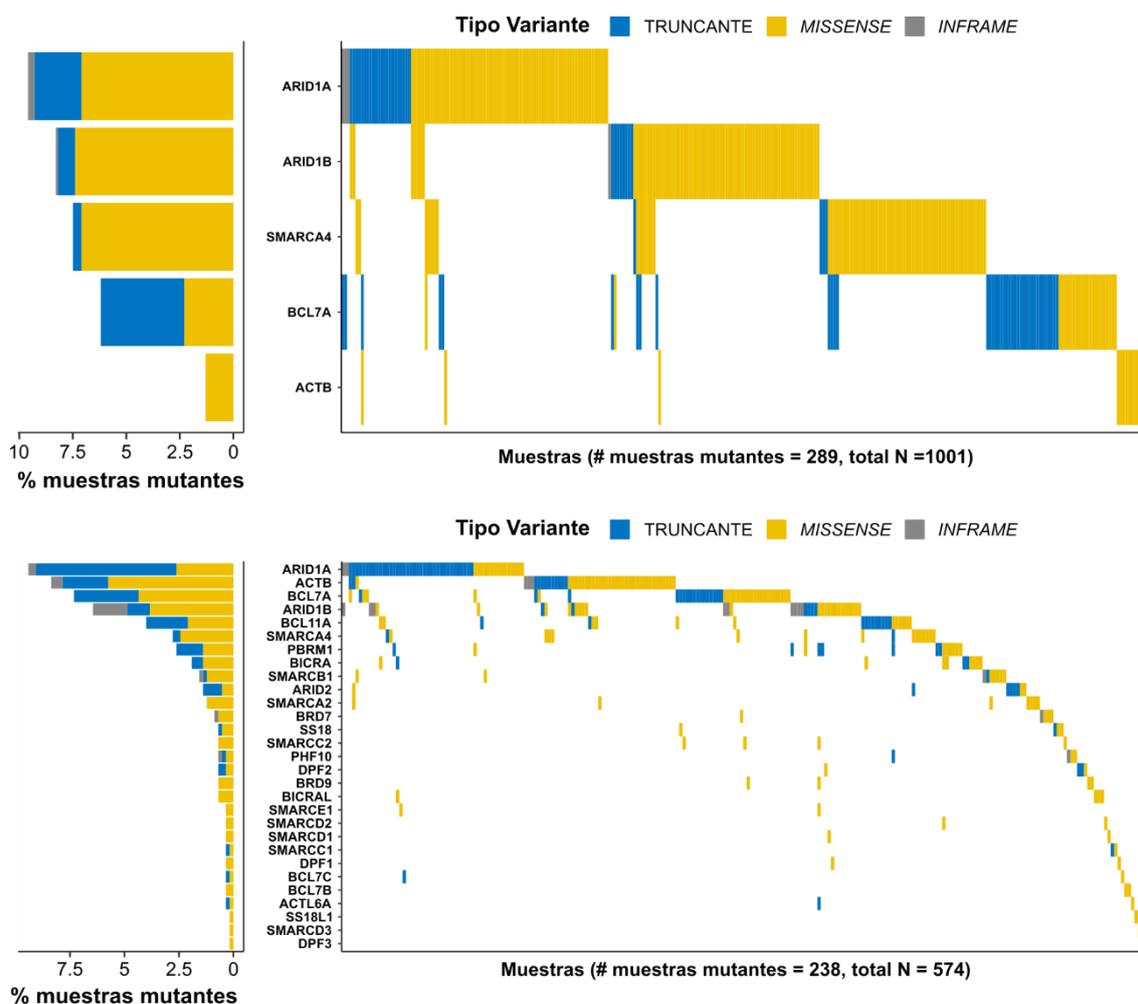


Figura 33. Waterfall plot de mutaciones en el complejo SWI/SNF en cohortes externas de DLBCL. Se muestran las mutaciones para el complejo SWI/SNF en las cohortes de Reddy et al. (panel superior) y Schmitz et al. (panel inferior). El eje X representa cada uno de los pacientes que tenía al menos una mutación en uno de los genes del complejo SWI/SNF. El número total de pacientes mutantes y el tamaño de la cohorte se muestran debajo del eje X. Si más de una mutación coexiste en el mismo gen para un mismo paciente, la mutación truncante se priorizó sobre otras mutaciones. Se consideraron como mutaciones truncantes a mutaciones en el sitio de *splicing*, mutaciones con cambio de marco de lectura (*frameshift*), mutaciones de pérdida del codón de inicio y mutaciones sin sentido. *Missense* = mutaciones con cambio de sentido. *IN-FRAME* = mutaciones con desplazamiento del marco de lectura. Los gráficos de barras representados a la izquierda muestran la frecuencia de mutación de cada uno de los genes del complejo SWI/SNF en las distintas cohortes.

Tabla 8. Análisis de enriquecimiento de mutaciones en genes SWI/SNF entre subtipos de DLBCL

Schmitz et al (2018), N = 574					
Gen	# mutaciones (N = 574)	# muts, GCB (N = 164)	# muts, ABC (N = 295)	# muts, No clasificado (NC) (N = 115)	Fisher p-valor
<i>ACTB</i>	48	20	18	10	0.080
<i>ACTL6A</i>	2	1	1	0	1
<i>ARID1A</i>	54	20	22	12	0.225
<i>ARID1B</i>	37	14	18	5	0.384
<i>ARID2</i>	8	1	5	2	0.639
<i>BCL11A</i>	23	7	12	4	1
<i>BCL7A</i>	42	22	13	7	0.003
<i>BCL7B</i>	2	0	2	0	0.706
<i>BCL7C</i>	2	0	1	1	0.442
<i>BICRA</i>	11	2	5	4	0.395
<i>BICRAL</i>	4	1	2	1	1
<i>BRD7</i>	5	0	4	1	0.411
<i>BRD9</i>	4	1	3	0	0.817
<i>DPF1</i>	2	1	1	0	1
<i>DPF2</i>	4	1	2	1	1
<i>DPF3</i>	1	0	1	0	1
<i>PBRM1</i>	15	4	5	6	0.173
<i>PHF10</i>	4	1	3	0	0.817
<i>SMARCA2</i>	7	1	3	3	0.354
<i>SMARCA4</i>	16	6	8	2	0.630
<i>SMARCB1</i>	9	2	7	0	0.259
<i>SMARCC1</i>	2	1	1	0	1
<i>SMARCC2</i>	4	1	2	1	1
<i>SMARCD1</i>	2	2	0	0	0.121
<i>SMARCD2</i>	2	0	1	1	0.442
<i>SMARCD3</i>	1	0	0	1	0.200
<i>SMARCE1</i>	2	0	1	1	0.442
<i>SS18</i>	4	1	3	0	0.817
<i>SS18L1</i>	1	0	0	1	0.200
<i>SWI/SNF Global</i>	238	85	108	45	0.006

Continúa en la siguiente página

Reddy et al (2017), N = 1001					
Gen	# mutaciones (N = 1001)	# muts, GCB (N = 331)	# muts, ABC (N = 313)	# muts, No clasificado (NC) (N = 131)	Fisher p-valor
<i>ACTB</i>	13	8	2	2	0.174
<i>ARID1A</i>	96	33	31	14	0.961
<i>ARID1B</i>	83	32	20	13	0.248
<i>BCL7A</i>	62	29	12	7	0.035
<i>SMARCA4</i>	76	31	18	12	0.184
SWI/SNF Global	289	113	75	40	0.017

Para cada una de las bases de datos, solamente se muestran los genes del complejo SWI/SNF que tenían al menos una mutación. Nota: En el estudio de Reddy et al. solamente 775 pacientes fueron clasificados por subtipo de linfoma GCB/ABC/NC en base a datos de RNA-Seq. La fila “SWI/SNF Global” indica el número de pacientes que presentan mutación en al menos un gen del complejo SWI/SNF. Para determinar el enriquecimiento de mutaciones entre subtipos de DLBCL se utilizó una prueba exacta de Fisher. Las celdas con p-valores significativos (no ajustados) están resaltadas en verde.

Posteriormente, se realizó un análisis de enriquecimiento de términos por ontología de genes utilizando la herramienta *enrichR*¹⁶⁵, y usando como input los 50 genes más mutados en la base de datos de Reddy et al¹³⁶. El propio *BCL7A* pasó a encontrarse en la lista de 50 genes más mutados al incluir tanto las mutaciones de *splicing* como otras mutaciones detectadas en nuestro estudio. El análisis a través de *enrichR* mostró un enriquecimiento significativo en vías relacionadas con la remodelación de la cromatina y genes SWI/SNF (**Tabla 9**). En general, nuestros resultados apuntan a un papel importante del complejo remodelador de cromatina SWI/SNF en DLBCL, especialmente en el subtipo GCB-DLBCL.

Hasta ahora, hemos visto como las mutaciones en el gen *BCL7A* se encuentran preferencialmente en el subtipo GCB-DLBCL, en concordancia con los análisis realizados en los propios estudios originales tanto de Reddy et al. como de Schmitz et al. Además, si nos fijamos con más detalle en la clasificación realizada por Schmitz et al., las mutaciones en *BCL7A* tienden a estar presentes, no solo en el subtipo GCB, sino más concretamente en el subtipo que definen como “EZB”. Este subtipo se caracteriza por presentar mutaciones en *EZH2* y translocaciones en *BCL2*.

Tabla 9. Términos GO enriquecidos al analizar los 50 genes más mutados en pacientes de DLBCL

Categoría GO	Términos GO	Puntuación combinada	p-valor ajustado	Genes
Cellular component	nuclear chromatin (GO:0000790)	31.2307779	1.3959E-05	CREBBP;MEF2B;MYC;HIST1H1E;ARID1A;TP53;EZH2;SMARCA4
Cellular component	SWI/SNF complex (GO:0016514)	27.0428651	0.00014583	ARID1A;ARID1B;SMARCA4
Cellular component	BAF-type complex (GO:0090544)	25.1532139	0.0001589	ARID1A;ARID1B;SMARCA4
Biological process	positive regulation of transcription, DNA-templated (GO:0045893)	35.91242	2.59E-07	CREBBP;MEF2B;TET2;ARID1A;ETS1;ARID1B;MTOR;DCAF6;SMARCA4;TBL1XR1;KLHL6;IRF4;MYC;IRF8;EP300;STAT6;TP53
Biological process	cytokine-mediated signaling pathway (GO:0019221)	32.99529	1.83E-08	CD70;PIK3CD;SMARCA4;SOCS1;IRF4;MYC;BCL2;PIM1;IRF8;TNFRSF14;STAT6;B2M;TP53;MYD88;MCL1
Biological process	chromatin-mediated maintenance of transcription (GO:0048096)	30.90866	0.00301	ARID1A;ARID1B
Molecular function	activating transcription factor binding (GO:0033613)	25.5298829	0.00045943	CREBBP;MEF2B;MYC;EP300
Molecular function	RNA polymerase II regulatory region sequence-specific DNA binding (GO:0000977)	24.5067747	0.00014351	MEF2B;IRF4;MYC;IRF8;ARID5B;STAT6;ETS1;TP53;SMARCA4
Molecular function	core promoter proximal region sequence-specific DNA binding (GO:0000987)	23.379652	6.5406E-05	CREBBP;MEF2B;IRF4;MYC;IRF8;ETS1;TP53;SMARCA4

Para cada una de las categorías de ontología génica (GO), se muestran los 3 términos con mayor puntuación. Para el análisis de enriquecimiento, mediante la herramienta enrichR, se utilizaron los 50 genes más mutados en la cohorte de Reddy et al., después de incluir las mutaciones en *BCL7A* detectadas en nuestro análisis. Los datos de enriquecimiento están infravalorados ya que *BCL7A* no está anotado en *Gene Ontology* como miembro del complejo SWI/SNF o como gen relacionado con la remodelación de la cromatina.

EZH2 es la subunidad catalítica del complejo represor Polycomb 2 (PRC2). La relación entre el complejo SWI/SNF y los complejos Polycomb tiene una gran importancia, debido a que tienen funciones opuestas tanto en células normales como en células tumorales¹⁶⁶. Por otro lado, se sabe que mutaciones concomitantes en SWI/SNF y *EZH2* tienen importantes implicaciones funcionales y terapéuticas^{167,168}. Además, en el contexto de los linfomas, se ha determinado que las mutaciones en *EZH2* están implicadas en el desarrollo tumoral y especialmente en el subtipo GCB-DLBCL^{169,170}. Teniendo esto en cuenta y considerando que nosotros identificamos nuevas mutaciones de *BCL7A*, decidimos analizar si las mutaciones en SWI/SNF ocurrían con mutaciones en los complejos PRC, con especial atención en la relación mutacional entre *BCL7A* y *EZH2*. Nuestro análisis mostró que había una coocurrencia significativa de mutaciones en *BCL7A* y *EZH2* en la cohorte de Schmitz et al. ($p = 0.009$; prueba exacta de Fisher) y una tendencia similar en la base de datos de Reddy et al., aunque no era significativa ($p = 0.09$). Sorprendentemente, las mutaciones en *BCL7A* solamente ocurrían con mutaciones *EZH2* en la tirosina 641 (*EZH2*^{Y641}) (**Tabla 10**). Las mutaciones *EZH2*^{Y641} son mutaciones de ganancia de función y se ha descrito que están presentes preferencialmente en el subtipo GCB-DLBCL¹⁶⁹. A excepción de *EZH2*, no se encontró ninguna asociación de coocurrencia o exclusión entre mutaciones en *BCL7A* y mutaciones en cualquier otra subunidad de los complejos PRC1 o PRC2 (resultados no mostrados).

Tabla 10. Coocurrencia de mutaciones de *BCL7A* y *EZH2* en cohortes externas

	Estudio de Schmitz et al.		Estudio de Reddy et al.	
	<i>BCL7A</i> silvestre	<i>BCL7A</i> mutante	<i>BCL7A</i> silvestre	<i>BCL7A</i> mutante
<i>EZH2</i> silvestre	489	33	885	55
<i>EZH2</i> mutante (No Y641)	10	0	4	0
<i>EZH2</i> ^{Y641}	33	9	50	7

Clasificación de pacientes de DLBCL en función a su estado mutacional de los genes *BCL7A* y *EZH2*. Los pacientes con más de una mutación en *EZH2* se clasificaron como “*EZH2* Y641” si una de las mutaciones afectaba a la posición Y641.

Posteriormente, se decidió ampliar el estudio a otras subunidades del complejo SWI/SNF, pero no encontramos ninguna relación de coocurrencia o exclusión de mutaciones entre las subunidades del complejo SWI/SNF y cualquiera de las subunidades de los complejos PRC1 o PRC2. Para cada una de las bases de datos, solo incluimos en el análisis las subunidades que presentaban mutaciones en al menos 10 pacientes.

A continuación, decidimos considerar el complejo SWI/SNF como un todo y clasificar las muestras como mutantes o silvestres para el complejo SWI/SNF en función de si tienen al menos una mutación en alguno de los genes que lo forman. En base a este análisis, se encontró una relación de coocurrencia significativa entre mutaciones en el complejo SWI/SNF y mutaciones en EZH2 en la base de datos de Schmitz ($p = 9,9E-5$) pero no en la base de datos de Reddy et al. ($p = 0,24$). Es importante tener en cuenta que, debido a motivos de filtrado, el estudio de Reddy et al. solamente reportó mutaciones en 5 de las 31 subunidades del complejo SWI/SNF (**Tabla 8**). Por otro lado, a diferencia de lo observado para *BCL7A*, el complejo SWI/SNF mutante (para alguna subunidad) no presentó coocurrencia con mutaciones EZH2^{Y641} de forma exclusiva, pero sí que conservó una tendencia significativa (**Tabla 11**), siendo pocos los casos en los que coexisten mutaciones en SWI/SNF y mutaciones en *EZH2* distintas a EZH2^{Y641}.

Tabla 11. Coocurrencia de mutaciones en el complejo SWI/SNF y en EZH2 en cohortes externas

	Estudio de Schmitz et al.		Estudio de Reddy et al.	
	SWI/SNF silvestre	SWI/SNF mutante	SWI/SNF silvestre	SWI/SNF mutante
EZH2 silvestre	387	135	672	268
EZH2 mutante (No Y641)	6	4	2	2
EZH2 ^{Y641}	18	24	37	20

Clasificación de pacientes de DLBCL en función a su estado mutacional en *EZH2* y en los genes del complejo SWI/SNF. Los pacientes con mutaciones en al menos un gen del complejo SWI/SNF fueron considerados como “SWI/SNF mutante”. Los pacientes con más de una mutación en *EZH2* se clasificaron como “EZH2 Y641” si una de las mutaciones afectaba a la posición Y641.

6. DISCUSIÓN

A mediados del año 2013, poco antes de comenzar esta tesis doctoral, *BCL7A* acababa de ser caracterizado como una “nueva” subunidad del complejo SWI/SNF en humanos⁶⁹. Además, mientras que algunos estudios de secuenciación masiva mostraban que *BCL7A* estaba mutado en determinadas neoplasias hematológicas, las mutaciones en *BCL7A* eran prácticamente nulas en otros tipos de tumores^{69,75}. Es por ello que pusimos en marcha un estudio piloto en busca de mutaciones o alteraciones en *BCL7A* en líneas celulares hematológicas. Este estudio nos condujo a profundizar en el DLBCL, siendo este el subtipo tumoral que aglomeraba la mayoría de variaciones genéticas encontradas. Al extender el estudio a un mayor número de muestras de DLBCL, se detectó una frecuencia mutacional del 17% sobre 41 líneas celulares y del 5.3% en una cohorte de 38 pacientes de DLBCL.

A diferencia de lo que nosotros observamos en nuestro estudio de secuenciación, nuevos trabajos que aplicaban NGS a tumores de DLBCL no definían a *BCL7A* como un gen frecuentemente mutado en este tipo de neoplasia^{46,47}. Ahora bien, estos primeros estudios sirvieron para poner de manifiesto la alta heterogeneidad del DLBCL tanto a nivel molecular como a nivel clínico. Esta alta heterogeneidad derivó incluso en discrepancias entre estudios de secuenciación a la hora de definir con exactitud qué genes están frecuentemente mutados en DLBCL⁴⁸. Al mismo tiempo, el bajo número de muestras utilizado en estos estudios pioneros dificultaba la caracterización de mutaciones que podrían ser fisiopatológicamente relevantes, simplemente por presentar frecuencias mutacionales relativamente bajas¹³⁵. La limitación del número de muestras es algo que nosotros mismos hemos experimentado con tan solo 38 pacientes de DLBCL en nuestra cohorte. Aun así, pudimos caracterizar una frecuencia mutacional en *BCL7A* del 5,3% mediante secuenciación de Sanger. Esta técnica nos permite un análisis visual directo de las mutaciones sin entrar en consideraciones estadísticas que aplican a grandes estudios de secuenciación masiva.

Más recientemente, a finales del año 2017 se publicó un trabajo que analizaba un total de 1001 pacientes de DLBCL (Reddy et al.¹³⁶). Unos meses más tarde, se publicó otro estudio independiente que analizaba 574 pacientes de DLBCL (Schmitz et al.¹³⁷). Ahora sí, ambos estudios mostraban que *BCL7A* se encuentra frecuentemente mutado en DLBCL. Sin embargo, pudimos observar que estos estudios habían excluido de sus análisis las mutaciones de *splicing* en la posición c.92+1. Es cierto que nosotros mismos no habíamos detectado la mutación de *splicing* en nuestra cohorte de 38 pacientes de DLBCL, pero sí que la habíamos identificado como la mutación más recurrente en el panel de 41 líneas celulares analizado. Además, pudimos detectar la mutación de *splicing* en, al menos, 1 de los 15 pacientes de FL analizados en nuestro estudio. Del mismo modo, un trabajo publicado en el año 2017, que analizó una cohorte de 105

pacientes de FL, definió a *BCL7A* como uno de los genes con mayor frecuencia mutacional y, en este caso, sí que habían tenido en cuenta la mutación de *splicing*¹³⁸. De hecho, la mutación de *splicing* representaba el 27,6% (8/29) de todas las mutaciones encontradas en el gen *BCL7A*.

En los estudios de análisis de secuenciación masiva en cohortes numerosas, es necesario aplicar filtros que simplifiquen la detección de mutaciones. De una forma u otra, los criterios de filtrado aplicados en los trabajos de Reddy et al. o Schmitz et al. han excluido las mutaciones de *splicing* de *BCL7A*, pudiendo infravalorarlas en comparación a otras mutaciones exónicas. Si bien las mutaciones en secuencias codificantes pueden tener un gran impacto sobre la función de la proteína, también hay numerosos casos en los que las mutaciones de *splicing* tienen un importante efecto deletéreo en la proteína resultante¹⁷¹. Hay distintas secuencias conservadas que influyen en el proceso de *splicing* y se localizan en posiciones intrónicas adyacentes a los exones o bien cerca de esas regiones¹⁷². Se ha llegado a estimar que un 10-15% de las enfermedades genéticas hereditarias están relacionadas con mutaciones en secuencias conservadas de *splicing*¹⁷³. La mayoría de las mutaciones afectan a los nucleótidos invariables AG o GT en los sitios de *splicing* 3' y 5' respectivamente¹⁷⁴. En el caso de *BCL7A*, la mutación en la guanina de la secuencia donadora del sitio de *splicing* 5' hace que el espliceosoma reconozca un sitio de *splicing* críptico en el propio exón uno, dando lugar a una delección de 27 aminoácidos. La formación de sitios crípticos de *splicing* es algo que se ha caracterizado en determinadas enfermedades hereditarias como son la beta talasemia¹⁷⁵. En el contexto del cáncer también se han caracterizado mutaciones particulares que generan un *splicing* alternativo con diferentes consecuencias moleculares¹⁷⁶. Igualmente, se han relacionado con cáncer patrones de *splicing* alternativo en determinados genes que no dependen de mutaciones puntuales en el propio gen sino en defectos en los componentes de la maquinaria de *splicing*¹⁷⁶. En conclusión, creemos que todo análisis mutacional debe tener en cuenta tanto mutaciones exónicas como mutaciones de *splicing*, ya que ambos tipos de mutaciones pueden tener importantes implicaciones funcionales.

Teniendo en cuenta la alta frecuencia de la mutación de *splicing* en las líneas celulares de DLBCL analizadas y en la cohorte de 105 pacientes de FL¹³⁸, es difícil asumir que en una cohorte combinada de 1575 pacientes de DLBCL no se hubiese descrito ninguna mutación de *splicing* en la posición c.92+1. En el estudio de Schmitz et al. sí que detectaron como somáticas las mutaciones en la posición de *splicing* c.92+2 pero, en cambio, no describieron ninguna mutación en la posición c.92+1. Esto se debe a que los autores originales descartaron las mutaciones presentes en bases de datos de población sana, tales como dbSNP, así como en colecciones internas no públicas. Las mutaciones en c.92+1 en *BCL7A* aparecen a frecuencias muy bajas en

bases de datos de población sana. En dbSNP, mutaciones en c.92+1 de *BCL7A* aparecen descritas con una frecuencia nula, lo cual es suficiente para que fuesen filtradas en estudios como el de Schmitz et al. En gnomAD, la mayor colección de exomas y genomas completos en población sana con más de 100.000 individuos, las mutaciones en posición c.92+1 en *BCL7A* aparecen con una frecuencia del orden de 10^{-5} , y en todos los casos son heterocigotas. Estas observaciones llevan a concluir que no hay motivos suficientes para excluir las mutaciones de *splicing* en la posición c.92+1 en *BCL7A* como irrelevantes funcionalmente en cáncer en base a su frecuencia en la población sana. Por otro lado, los autores del estudio de Reddy et al. directamente restringieron el análisis a mutaciones exónicas, obviando cualquier posición de *splicing* a lo largo de todo el genoma.

Al analizar nosotros mismos las dos cohortes externas de DLBCL e incluir las posiciones de *splicing* en el análisis, se encontró un incremento del 27,4% y del 17% en las mutaciones totales de *BCL7A* en los estudios de Reddy et al. y Schmitz et al. respectivamente. Ahora bien, nuestro análisis tenía una limitación para poder determinar el estado somático de las mutaciones de *splicing* caracterizadas ya que no disponíamos de datos de secuenciación de línea germinal o tejido sano adyacente al tumor biopsiado. Aun así, hemos mostrado evidencia suficiente en favor de que las mutaciones de *splicing* no están presentes en línea germinal. Además, hemos caracterizado cómo las mutaciones de *splicing*, cuando coocurren con otras mutaciones inactivadoras en *BCL7A*, presentan un patrón de inactivación bialélica. La inactivación bialélica es una característica que se encuentra frecuentemente en genes supresores de tumores¹⁷⁷. Tras el estudio pionero de Knudson en 1971⁸, la detección de un patrón de inactivación bialélica en determinados genes ha servido como aliciente para iniciar diferentes estudios funcionales que han demostrado el papel supresor tumoral de los genes inactivados. Además de coocurrir en ambos alelos, muchas de las mutaciones inactivadoras presentaban valores de frecuencia alélica elevados, lo cual es característico de mutaciones que se generan en etapas tempranas del desarrollo tumoral. Globalmente, el gen *BCL7A* muestra patrones mutacionales típicos de los genes supresores de tumores en el contexto del DLBCL, donde podría ejercer un rol importante desde etapas tempranas del desarrollo tumoral.

Otro detalle importante acerca del patrón mutacional de *BCL7A* es que las mutaciones se agrupan en el dominio amino-terminal, más concretamente en el exón 1. Es por ello que su etiología se ha ligado al mecanismo de SHM^{137,144}. Anteriores trabajos se basaban en combinar distintos parámetros para estimar grupos de genes que sufren el proceso de SHM. Estos parámetros incluyen la localización o tipo de mutación observada, el nucleótido concreto afectado o los niveles de expresión génica¹⁴⁴. Aun así, estos parámetros simplemente predicen

la acción de la enzima AID. En cambio, nosotros hemos aportado evidencia experimental de como *BCL7A* sufre mutagénesis dirigida por la acción de la enzima AID utilizando modelos murinos diseñados para este fin. De la misma forma, la acción directa de la enzima AID sobre *BCL7A* también ha sido demostrada previamente en estudios *in vitro*¹⁴³. Por otra parte, la asociación de las mutaciones de *BCL7A* con el proceso de SHM no solo se ha descrito en DLBCL sino también en otros linfomas derivados del CG. Así, mutaciones en *BCL7A* tanto en regiones codificantes como no codificantes son más frecuentes en el subtipo endémico de linfoma de Burkitt que en los subtipos esporádico o el asociado a inmunodeficiencia. Es precisamente el subtipo endémico el que presenta una mayor actividad global de la enzima AID¹⁷⁸. Por otro lado, mutaciones derivadas de la acción de AID en *BCL7A* se han relacionado con la transformación de FL a DLBCL¹⁵². Además, la huella de AID podría también explicar las mutaciones de *BCL7A* en linfomas derivados de células B post-CG como el mieloma múltiple, donde mutaciones en *BCL7A* se han identificado en regiones no codificantes cercanas al inicio de la transcripción¹³⁴.

Diferentes evidencias muestran que *BCL7A* sufre mutaciones mediante el proceso de SHM. Dado que los mecanismos de SHM siguen activos en el subtipo GCB, pero no en el subtipo ABC³⁷, esto podría explicar el hecho de que las mutaciones de *BCL7A* aparezcan con mayor frecuencia en el subtipo GCB en comparación con el subtipo ABC (Ver artículos originales de Reddy y Schmitz). También, se podría razonar que las mutaciones en *BCL7A* son una mera consecuencia del contexto tumoral en el que nos encontramos. De hecho, siempre hay grandes dudas en la comunidad científica sobre si las mutaciones derivadas de AID son mutaciones pasajeras o realmente podrían representar mutaciones conductoras con capacidad de promover el desarrollo tumoral. Para intentar resolver esta dicotomía, lo primero que hay que tener en cuenta es que muchas de las mutaciones generadas por AID, en genes distintos a los de inmunoglobulinas, son reparadas cooperativamente por distintos sistemas de reparación^{148,179}. Aun así, algunas mutaciones sin reparar podrían proporcionar una ventaja proliferativa. De esta forma, en caso de ser seleccionadas positivamente, determinadas mutaciones inducidas por AID podrían detectarse frecuentemente en determinados linfomas. En células B normales no suelen detectarse mutaciones originadas por SHM fuera de los locus de genes de inmunoglobulinas, aunque se han demostrado determinadas excepciones como ocurre con el gen *BCL6*^{180,181}. En consecuencia, se ha descrito que determinadas mutaciones originadas mediante SHM en *BCL6* aparecen únicamente en pacientes con DLBCL, lo cual sugiere un mecanismo de selección positivo por parte de las células tumorales²⁹. Además de *BCL6*, otras mutaciones encontradas frecuentemente en linfomas derivados del CG en genes como *PIM1* o *GNA13* se han

caracterizado siguiendo la metodología empleada en nuestro estudio¹⁴⁸, lo que sugiere que el mecanismo de SHM puede ser el origen de ciertas mutaciones con potencial linfomagénico.

En el caso de *BCL7A*, se desconoce si este gen presenta mutaciones originadas por SHM en células B normales. Ahora bien, hemos podido determinar que algunas de las mutaciones de *BCL7A* en la región amino-terminal pueden originarse mediante la acción de AID. Aun así, en base a nuestros resultados, podemos concluir que no hay una relación directa que implique que una mutación originada por AID tenga una mayor prevalencia en DLBCL. Algunas de las citosinas o guaninas dentro de secuencias de reconocimiento de AID (RGYW/WRCY), que aparecen mutadas en nuestros modelos murinos, ni siquiera aparecen mutadas en los 1575 pacientes de DLBCL analizados (Figura 26). En cuanto a la mutación de *splicing*, aunque no presenta una frecuencia de transición alta en los modelos murinos, se encuentra dentro de un motivo RGYW y resulta ser la mutación más frecuente de *BCL7A* en pacientes de DLBCL. En cuanto a otras posiciones del gen *BCL7A* que representan puntos calientes mutacionales en pacientes de DLBCL, como el codón de inicio o el codón R11, no se encuentran bajo secuencias de reconocimiento por AID y no son susceptibles de mutar mediante SHM. En conclusión, el proceso de SHM podría originar las mutaciones inactivadoras más prevalentes en pacientes de DLBCL, las mutaciones de *splicing*, pero la acción de la enzima AID no es el único mecanismo por el cual se originan mutaciones inactivadoras en *BCL7A*. En conjunto a la acción de la enzima AID, la alta prevalencia de la mutación de *splicing* sugiere un mecanismo adicional de selección por parte de las células tumorales para adquirir una ventaja proliferativa.

6.1 LAS MUTACIONES DE *SPLICING* DE *BCL7A* IMPIDEN SU UNIÓN AL COMPLEJO SWI/SNF

Nuestro estudio ha puesto de manifiesto que la mutación de *splicing* es la más frecuente de las mutaciones detectadas en *BCL7A* en pacientes de DLBCL. Más allá de su alta prevalencia, hemos demostrado que las mutaciones de *splicing* dan lugar a una proteína *BCL7A* truncada ($\Delta 27$ -*BCL7A*), que pierde 27 aminoácidos de su región amino-terminal. La región perdida forma parte de los 51 aminoácidos que componen el dominio BCL_N (Anotado en la base de datos Pfam: "PF04714- *BCL*, *N-terminal conserved region*"), el cual se comparte entre los genes parálogos *BCL7A*, *BCL7B* y *BCL7C*¹²⁷. Nosotros hemos nombrado "dominio BCL7_Nt" al fragmento de 27 aminoácidos que se pierde debido a las mutaciones de *splicing*. Los genes parálogos *BCL7A*, *BCL7B* y *BCL7C*, de ahora en adelante nombrados de forma conjunta como genes o subunidades BCL7, son subunidades excluyentes a la hora de formar los complejos SWI/SNF. El hecho de que compartan el dominio BCL_N sugiere una función común para dicho dominio en relación a la unión de las diferentes subunidades BCL7 al complejo SWI/SNF. Es por ello que decidimos

realizar experimentos de inmunoprecipitación para evaluar las consecuencias de las mutaciones de *splicing* sobre la unión de BCL7A con el complejo SWI/SNF. De tal forma, nuestros análisis revelaron que la pérdida del dominio BCL7_Nt hace que la proteína resultante $\Delta 27$ -BCL7A no sea capaz de unirse al resto del complejo SWI/SNF. Absolutamente todas las subunidades del complejo SWI/SNF que inmunoprecipitan con la forma silvestre de BCL7A dejan de hacerlo con el mutante $\Delta 27$ -BCL7A. Por lo tanto, podemos concluir que el dominio BCL7_Nt es totalmente necesario para que BCL7A pueda unirse a cualquier subunidad del complejo SWI/SNF, mientras que el resto de la secuencia peptídica de BCL7A no realizaría ninguna función en este aspecto en base a nuestros resultados.

Recientemente, se ha descrito que la subunidad BCL7A y sus homólogos BCL7B y BCL7C forman parte del denominado módulo ATPasa dentro del complejo SWI/SNF⁷⁷. Este módulo está formado por la subunidad catalítica excluyente SMARCA4 o SMARCA2 junto a las subunidades excluyentes ACTL6A o ACTL6B y la subunidad ACTB. El complejo ATPasa también lo forma la subunidad SS18, la cual no está presente en los complejos PBAF. En cambio, las subunidades BCL7 sí que se ha descrito que están presentes en el módulo ATPasa de forma ubicua en los diferentes complejos cBAF, PBAF o ncBAF⁷⁷. A la hora de ensamblarse el complejo SWI/SNF, el módulo ATPasa parece ser el último en unirse en bloque al resto del complejo, que ya está ensamblado previamente (Ver figura 4). Esto concuerda con lo que hemos observado nosotros mismos en cuanto a que el estado de BCL7A, silvestre o mutante, no afecta a la integridad de parte del complejo residual ya que la subunidad SMARCA4 mantiene sus uniones con las otras subunidades que forman el núcleo principal de los complejos BAF. Ahora bien, en futuros estudios y dado que BCL7A pertenece al módulo ATPasa, sería interesante comprobar si la pérdida del dominio BCL7_Nt podría afectar también a la unión de los otros componentes dentro del propio módulo ATPasa, así como determinar cuál podría ser la función de BCL7A dentro del módulo ATPasa. En el caso de las subunidades ACTL6 o ACTB (β -actina) sí que se ha determinado que son necesarias para potenciar la actividad ATPasa de SMARCA4¹⁸². También sería interesante comprobar si al perder BCL7A su asociación con el complejo SWI/SNF, cabría la posibilidad de que su lugar fuese ocupado por las subunidades homólogas BCL7B o BCL7C. Por el momento, debido a la ausencia de anticuerpos comerciales que den buenos resultados para revelar estas proteínas mediante Western blot, no pudimos responder a esta cuestión.

Hace apenas unos meses, se ha publicado un estudio que puede ser de gran importancia para descifrar la función que podría tener BCL7A dentro del complejo SWI/SNF y más concretamente dentro del módulo ATPasa. Mediante la combinación de técnicas innovadoras como cryo-EM y CX-MS (del inglés "*cryoelectron microscopy*" y "*cross-linking mass spectrometry*"), se han

estudiado los complejos BAF a nivel estructural, incluso en su conformación de unión a nucleosomas^{183,184}. Gracias a estos estudios, se ha conseguido por primera vez mapear la posición relativa de las proteínas BCL7 dentro del complejo. Además, se han caracterizado interacciones específicas de BCL7A con otras subunidades del complejo. En este sentido, se ha visto que el dominio amino-terminal de BCL7A (BCL_N) se une al dominio HSA de SMARCA4/2. El dominio HSA es capaz de unir actina o bien proteínas ARP relacionadas con actina (ARP; *Actin-related proteins*) como son ACTB o ACTL6. Precisamente, son las proteínas ARP las que forman parte del módulo ATPasa junto a BCL7A. Además, BCL7A también se une a determinados residuos de la proteína SMARCB1 y de las histonas nucleosomales H2A y H2B¹⁸³. Por tanto, debido a su localización dentro del módulo ATPasa y el lugar que ocupa dentro del complejo, BCL7A podría tener una implicación dentro de la actividad remodeladora de la cromatina del complejo SWI/SNF en conjunto a otras proteínas relacionadas con la actina. Por último, mientras que el dominio BCL7_Nt es crítico para unirse al complejo SWI/SNF, la función del resto de la secuencia peptídica de BCL7A se desconoce. Futuros estudios podrán determinar si BCL7A puede aportar al complejo SWI/SNF funciones concretas aún por dilucidar.

6.2 LA RESTAURACIÓN DE LA EXPRESIÓN DE BCL7A EJERCE UN PAPEL SUPRESOR TUMORAL EN MODELOS CELULARES DE DLBCL

Uno de los objetivos de este estudio era definir líneas celulares que pudiesen ser usadas como modelo para estudiar la actividad de BCL7A. En un principio, se pensó que los modelos celulares óptimos serían aquellos en los que se hubiese perdido la expresión de BCL7A. Consecuentemente, las líneas celulares deficientes en BCL7A podrían ser usadas para restaurar la expresión de este gen y evaluar las implicaciones funcionales de BCL7A sobre el fenotipo tumoral. Ahora bien, nuestro estudio nos ha dirigido hacia un contexto diferente donde no existe una pérdida de expresión de BCL7A como tal. Después de un exhaustivo análisis en 41 líneas celulares de DLBCL, nos encontramos con modelos celulares que no expresaban una forma salvaje de BCL7A, pero sí que expresaban distintas formas mutantes.

Los estudios funcionales realizados en este trabajo se han llevado a cabo en líneas celulares donde ambos alelos de *BCL7A* presentaban diferentes eventos genéticos, que incluyen mutaciones y duplicaciones. En cuanto a las mutaciones encontradas, se han definido mutaciones puntuales en el codón de inicio de la traducción, mutaciones con cambio de sentido y mutaciones de *splicing*. Estas últimas, detectadas en las líneas celulares Karpas-1106, OCI-LY1 y VAL, dan lugar a una delección del dominio BCL7_Nt. Por otro lado, en la línea celular Karpas-1106 se caracterizó una duplicación en tándem de los exones 2 y 3. Mediante análisis *in silico*,

se comprobó que la duplicación generaba un codón de terminación prematuro y se descarta que pueda generarse una proteína BCL7A aberrante a partir de ese alelo. Mutaciones con cambio de sentido se detectaron en las líneas celulares DB, OCI-LY7, ROS-50 o VAL (Apéndice II). En el caso de DB y OCI-LY7, estas líneas celulares siguen expresando un alelo salvaje de BCL7A por lo que fueron descartadas para usarse como modelos. La línea celular ROS-50 presentaba mutaciones heterocigotas con cambio de sentido en ambos alelos. Dada la baja frecuencia de estas mutaciones en la cohorte de DLBCL analizada, se descartó el uso de ROS-50 en los estudios funcionales, ya que no hay evidencia de que esas mutaciones afecten a la función de BCL7A.

Por tanto, las líneas celulares candidatas a usarse en los estudios funcionales fueron Karpas-1106, OCI-LY1 y VAL. Por un lado, tenemos las líneas celulares Karpas-1106 y OCI-LY1 que expresan el mutante $\Delta 27$ -BCL7A. Por otro lado, tenemos la línea celular VAL que expresa tanto el mutante $\Delta 27$ -BCL7A como un BCL7A con mutación en el codón 11 (R11T-BCL7A). La posición del codón R11, a diferencia de otras mutaciones con cambio de sentido, sí que representa un punto caliente mutacional y podría ser de gran valor estudiar un fondo genético donde coocurren la mutación en R11 y la mutación de *splicing*. Una vez caracterizado el estatus de BCL7A en estos tres modelos celulares mencionados, se procedió a restaurar la expresión de la forma salvaje de BCL7A mediante el uso de sistemas de transducción lentiviral. Además, también usamos como control las transducciones de la forma mutante $\Delta 27$ -BCL7A y de un vector vacío. La línea celular Karpas-1106 resultó ser un modelo celular muy complicado con el que trabajar ya que las eficiencias de transducción eran tan bajas que no pudimos aislar por FACS poblaciones celulares viables que expresasen nuestras construcciones. Es por ello que los estudios funcionales se limitaron a los modelos celulares de OCI-LY1 y VAL.

La restauración de la expresión de BCL7A en las líneas celulares OCI-LY1 y VAL dio lugar a una disminución de la proliferación. Estos cambios fenotípicos fueron determinados gracias al desarrollo de ensayos de crecimiento celular competitivo. Este tipo de ensayos se adaptan bien al uso de líneas celulares hematológicas^{121,122}. Al mismo tiempo, estos ensayos también encajan con la metodología que hemos usado, a través de la cual conseguimos la expresión estable del transgén a lo largo del tiempo mediante transducción lentiviral. Por tanto, los ensayos de crecimiento celular competitivo permiten la detección de cambios sutiles en la proliferación celular, pero prolongados en el tiempo¹⁵¹. De tal forma, pudimos identificar una disminución en la proliferación en las células que expresan BCL7A, mientras que este fenotipo no se observa al expresar la forma mutante $\Delta 27$ -BCL7A o el vector vacío. Posteriormente, usamos la línea celular OCI-LY1 para llevar a cabo xenotrasplantes en ratones NSG y analizar *in vivo* el crecimiento de la línea tumoral trasplantada. De esta forma, pudimos corroborar el fenotipo previamente

observado en cultivos celulares. Por tanto, podemos concluir que restaurar la expresión de BCL7A ejerce un efecto supresor tumoral en modelos celulares de DLBCL que expresan la forma mutante $\Delta 27$ -BCL7A.

Aunque en este trabajo nos hemos centrado en caracterizar las consecuencias moleculares y funcionales de las mutaciones de *splicing* en BCL7A, es probable que las mutaciones en R11 puedan tener también un impacto importante en la función de BCL7A. Estas mutaciones a menudo coexisten con mutaciones inactivadoras y, cuando esto ocurre, tienden a estar presentes en diferentes alelos. Además de presentar un patrón de “inactivación” bialélica, la mutación en R11 es la única mutación con cambio de sentido que ha sido caracterizada como punto caliente mutacional. En conjunto, estas apreciaciones ponen de manifiesto la importancia de evaluar si las mutaciones en R11 afectan y cómo a la función de BCL7A. Por el momento, en función a los ensayos realizados en la línea celular VAL, no podemos concluir si el efecto anti-proliferativo se debe a rescatar la función del alelo mutante $\Delta 27$ -BCL7A o también se restaura la función “perdida” por la mutación en la posición R11.

Una vez caracterizado el fenotipo supresor tumoral que ejerce el rescatar la expresión de BCL7A, pasamos a estudiar posibles cambios en los perfiles de expresión génica. Los análisis de expresión diferencial mediante RNA-Seq revelaron algunos genes que pueden estar detrás de esa disminución proliferativa como pueden ser *TP63*, *TRIB2*, o *CDKN1A* (p21). Por un lado, el gen supresor tumoral *TP63*, el cual se ha visto que modula la transcripción de *CDKN1A*¹⁶², se mostró inducido en células que expresan BCL7A silvestre. Por otro lado, encontramos una inhibición en la expresión de *TRIB2*, que actúa como oncogén en neoplasias hematológicas¹⁶³ y que recientemente se ha caracterizado que es capaz de promover la tumorigénesis mediante el bloqueo de la vía de señalización AP4/*CDKN1A*¹⁶⁴. Desconocemos si puede haber una relación directa entre *TP63* y *TRIB2* con la transcripción del gen *CDKN1A*, pero este último mostró mayores niveles de expresión, incluso a nivel de proteína, en células que expresan BCL7A silvestre.

Por otro lado, los estudios de expresión diferencial han mostrado un claro enriquecimiento en vías de activación de células B, lo cual puede tener gran relevancia en el contexto del DLBCL. Mientras que hay estudios que demuestran la importancia de un gen como *SMARCA4* en establecer el linaje de una célula B en estadios tempranos¹⁸⁵, a día de hoy se desconoce en gran parte las funciones que puede ejercer el complejo SWI/SNF durante la reacción de CG. Futuros estudios permitirán comprender las implicaciones de BCL7A y del complejo SWI/SNF en el contexto del CG. El abordaje óptimo de estos estudios debería realizarse en modelos *in vivo*,

donde las células B, dentro de su nicho biológico, puedan interactuar con otros tipos celulares como linfocitos T o células foliculares dendríticas a lo largo del proceso de maduración por afinidad.

6.3 EL COMPLEJO SWI/SNF EN EL CONTEXTO FISIOPATOLÓGICO DEL CENTRO GERMINAL

El DLBCL es una expansión maligna de clones de células B que han quedado atrapados en alguna de las diferentes etapas de diferenciación de una célula B a lo largo de la reacción de CG. Los diferentes pasos que va dando una célula B durante el proceso de maduración por afinidad están controlados a través de una regulación transcripcional muy precisa. Es por ello que diferentes lesiones génicas que dañan gran variedad de genes involucrados en la regulación transcripcional están relacionadas con la linfomagénesis¹⁷. Al mismo tiempo, la regulación transcripcional depende en gran medida de la estructura de la cromatina y de los diferentes mecanismos epigenéticos que la regulan¹⁸⁶. De forma conjunta, son muchos los genes que contribuyen a modificar covalentemente el ADN o las histonas para regular el estado conformacional de la cromatina entre una forma activa o represiva¹⁸⁷. En este sentido, los estudios de secuenciación masiva en DLBCL han mostrado que hay una gran cantidad de genes relacionados con la modificación de la cromatina que están mutados en DLBCL y otros linfomas derivados del CG. Estos genes incluyen metiltransferasas como *KMT2D (MLL2)* o acetiltransferasas como *CREBBP* o *EP300*¹⁸⁸. Todo ello apunta a un papel muy relevante de la regulación epigenética tanto en la fisiología como en la fisiopatología de los linfomas derivados del CG.

A día de hoy, se conoce muy poco acerca de la función que ejerce el complejo SWI/SNF en el contexto del CG. Algún estudio ha mostrado que SMARCA4 se requiere para regular la expresión de algunos genes relacionados con la proliferación de células B activadas¹⁸⁹, pero no hay indicios de que el complejo SWI/SNF esté relacionado con el proceso de linfomagénesis. Ahora bien, el complejo SWI/SNF podría tener una implicación muy importante en prevenir la linfomagénesis debido al rol que ejerce en antagonizar la función de los complejos represores Polycomb (PRC)¹⁶⁶. Investigaciones recientes en tumores sólidos han demostrado que la pérdida de la integridad del complejo SWI/SNF afecta a sus capacidades para desplazar a los complejos PRC de la cromatina¹⁹⁰, actividad para la cual es imprescindible la función ATPasa del complejo SWI/SNF¹⁹¹. En este sentido, se sabe que las células B, para entrar transitoriamente en el proceso de maduración, necesitan inhibir genes como *PRDM1* o *IRF4* para bloquear su diferenciación a células plasmáticas. Esta función es llevada a cabo de forma cooperativa por genes como *BCL6* o *EZH2*¹⁹², que es la subunidad catalítica del complejo PRC2. Además, también se inhiben genes como *CDKN1A* o *CDKN1B* con el objetivo de permitir un estado proliferativo más permisivo¹⁹³.

Todos estos genes, que se reprimen de forma transitoria, tienen en común que están bajo el control de promotores bivalentes¹⁹². Es precisamente en promotores bivalentes donde se ha visto que el complejo SWI/SNF es importante para contrarrestar la represión que ejercen los complejos PRC⁹³. En resumen, en el contexto de la linfomagénesis derivada del CG, un complejo SWI/SNF aberrante podría no ser capaz de desplazar los complejos PRC. Esta situación podría inducir un estado represivo exacerbado a nivel transcripcional dictado por los complejos PRC, quedando las células B bloqueadas en su diferenciación hacia células plasmáticas.

En los modelos celulares utilizados en nuestro estudio, pudimos observar que la expresión del gen *CDKN1A* (p21) fue inducida al restaurar la expresión de *BCL7A*. Este efecto se pudo ver mejor en la línea celular OCI-LY1, donde observamos un aumento de los niveles de p21 incluso a nivel de proteína. Esto podría estar relacionado con que, debido a la restauración de *BCL7A*, el complejo SWI/SNF recupera en parte su capacidad de antagonizar la represión de los complejos PRC sobre la transcripción del gen *CDKN1A*. De todas formas, estas observaciones no son suficientes para afirmar que el complejo SWI/SNF antagoniza a los complejos PRC en células B. Por el contrario, hay un estudio que afirma más bien lo contrario. Aunque no relacionaron el complejo SWI/SNF directamente con los complejos PRC, Choi, J. et al. definieron que el complejo SWI/SNF regula la formación del CG mediante la inhibición, dependiente de *BCL6*, de la expresión de *PRDM1*¹⁹⁴. Según este estudio, el complejo SWI/SNF sería necesario para ayudar en la represión transcripcional producida durante el proceso de maduración por afinidad de una célula B, lo cual podría ser incompatible con la función del complejo SWI/SNF en antagonizar a los complejos represores PRC. Aun así, ambos escenarios opuestos serían compatibles en base a nuevos estudios que indican que el complejo SWI/SNF es capaz tanto de promover la activación como la represión génica redistribuyendo la cantidad/dosis de los complejos represivos PRC de forma dinámica¹⁹⁵. Por otro lado, también se ha visto que otros complejos remodeladores de la cromatina como son los complejos Mi-2/NuRD son capaces de interactuar y cooperar con represores transcripcionales como *BCL6* o *EZH2*^{196,197}. Curiosamente, nosotros también hemos observado en nuestros análisis de inmunoprecipitación cómo el complejo SWI/SNF interactúa con miembros de los complejos Mi-2/NuRD (Apéndice IV). En conclusión, diferentes complejos represores y activadores parecen interactuar para dirigir una compleja regulación transcripcional y futuros estudios podrán definir las aportaciones individuales de cada uno de los participantes.

Tal y como hemos descrito previamente, hay multitud de genes implicados en la biología del CG que se encuentran mutados en linfomas. Así, se conoce que mutaciones de ganancia de función en *EZH2*, o bien una expresión constitutiva de *BCL6*, dan lugar a una hiperplasia del CG^{170,198}. De

hecho, las mutaciones de ganancia de función de *EZH2* (*EZH2*^{Y641}) son un sello distintivo de los linfomas derivados del CG como el FL o el subtipo GCB-DLBCL¹⁶⁹. Aun así, un *EZH2* mutante no parece ser suficiente para la transformación de células B en células tumorales. En cambio, mutaciones *EZH2*^{Y641} en combinación con la sobreexpresión de oncogenes como *Bcl2* o *Myc* aceleran la aparición de linfomas en modelos murinos^{170,199}. Del mismo modo, la inactivación del complejo SWI/SNF y las mutaciones de ganancia de función de *EZH2*^{Y641} podrían cooperar para limitar la diferenciación de las células B a células plasmáticas y, por tanto, inducir la linfomagénesis. De hecho, las mutaciones en *BCL7A* y *EZH2* tienden a coexistir en GCB-DLBCL, y más precisamente en el subtipo "EZB" descrito por Schmitz et al.¹³⁷, caracterizado por presentar mutaciones en *EZH2* y translocaciones en *BCL2*. Curiosamente, hemos visto que las mutaciones en *BCL7A* coocurren con mutaciones *EZH2*^{Y641} en las cohortes analizadas (**Tabla 10**). Estas observaciones podrían ser de gran importancia a la hora de la estratificar pacientes de DLBCL que puedan beneficiarse de determinadas terapias epigenéticas dirigidas.

Recientemente, diferentes estudios han mostrado que los tumores mutantes para el complejo SWI/SNF dependen de la actividad de *EZH2*²⁰⁰. Además, la inhibición de *EZH2* ha mostrado que puede ser una herramienta terapéutica prometedora contra los tumores mutantes en SWI/SNF¹⁶⁷. A principios del año 2020, el inhibidor de *EZH2* *Tazemetostat* fue aprobado por la FDA (*Food and Drug Administration*) para tratar sarcomas epitelioides carentes de la expresión de *SMARCB1*. En el contexto de pacientes con linfoma, hay evidencias preclínicas en las que inhibidores de *EZH2* junto a glucocorticoides han mostrado un efecto antitumoral en modelos de linfomas derivados del CG²⁰¹. Además, hay ensayos clínicos en curso para tratar con *Tazemetostat* a pacientes con linfoma no Hodgkin recidivante que tienen mutaciones de ganancia de función en *EZH2* y/o mutaciones de pérdida de función en *SMARCA4* o *SMARCB1*²⁰². Dada la alta frecuencia mutacional de otras subunidades del complejo SWI/SNF como *BCL7A* en linfomas derivados del CG, sería interesante incluir a pacientes mutantes para *BCL7A* y otras subunidades del complejo SWI/SNF en futuros ensayos clínicos que prueben inhibidores de *EZH2* en pacientes con linfomas.

7. CONCLUSIONES

- El gen *BCL7A* está mutado frecuentemente en muestras de linfoma difuso de células B grandes (DLBCL), mostrando una frecuencia mutacional del 17% en líneas celulares (7/41) y del 5,3% en una cohorte de pacientes (2/38).
- Las mutaciones de *BCL7A* en DLBCL se concentran en torno al exón 1. Además, *BCL7A* presenta tres puntos calientes mutacionales situados en el primer codón de la traducción, en el codón 11 y en la secuencia donadora de *splicing* del intrón 1, siendo este último el más frecuente.
- El reanálisis mutacional de una cohorte externa de 1575 pacientes de DLBCL desveló que se habían excluido, al menos parcialmente, las mutaciones del sitio donador de *splicing* del intrón 1. Su inclusión dio lugar a un incremento del 30% en la frecuencia mutacional. Así, la prevalencia mutacional de *BCL7A* en un total de 1575 pacientes de DLBCL fue del 6,6%.
- Mutaciones inactivadoras en *BCL7A*, junto a mutaciones en el codón 11, presenta un patrón de inactivación bialélica en muestras de DLBCL, una característica típica de genes supresores de tumores. En un 87% de los pacientes que mostraron más de una mutación inactivadora o mutación en el codón 11, estas mutaciones se encontraron en diferentes alelos.
- La mutación en el sitio donador de *splicing* genera un transcrito aberrante en el que, debido a un *splicing* críptico, se produce la pérdida de 81 nucleótidos del exón 1. Esto da lugar a la expresión de una proteína mutante que pierde 27 aminoácidos de su dominio amino-terminal ($\Delta 27$ -*BCL7A*).
- La pérdida de los 27 aminoácidos del dominio amino-terminal (*BCL7_Nt*) impide la unión de *BCL7A* con el resto de subunidades que forman el complejo SWI/SNF. En cambio, la pérdida del dominio *BCL7_Nt* no afecta a la unión del resto de subunidades centrales del complejo SWI/SNF.
- La restauración de la expresión de *BCL7A* en líneas celulares que expresan la forma mutante $\Delta 27$ -*BCL7A* da lugar a una disminución de la proliferación, lo cual es indicativo de su función supresora tumoral en este contexto tanto en cultivo celular como *in vivo*.

- La restauración de la expresión de BCL7A está relacionada con la regulación transcripcional de genes implicados en la activación de células B, así como otros genes relacionados con la proliferación y el ciclo celular.
- Tanto BCL7A como el complejo SWI/SNF están mutados con mayor frecuencia en el subtipo GCB-DLBCL en comparación con el subtipo ABC-DLBCL, lo cual puede tener implicaciones importantes en la estratificación de pacientes para la aplicación de determinadas terapias dirigidas.

8. BIBLIOGRAFÍA

1. Massague J, Obenauf AC. Metastatic colonization by circulating tumour cells. *Nature*. 2016;529(7586):298-306.
2. Stewart BW, Wild C, International Agency for Research on Cancer, World Health Organization. *World cancer report 2014*. Lyon, France
3. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70.
4. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.
5. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012;481(7381):306-313.
6. Croce CM. Oncogenes and cancer. *N Engl J Med*. 2008;358(5):502-511.
7. Weinberg RA. Tumor suppressor genes. *Science*. 1991;254(5035):1138-1146.
8. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A*. 1971;68(4):820-823.
9. Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. *Mol Cell*. 2010;40(2):179-204.
10. Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Lyon: International Agency for Research on Cancer (IARC); 2016
11. Horner MJ, Ries LAG, Krapcho M, Neyman N, Aminou R, Howlader N, Altekruse SF, Feuer EJ, Huang L, Mariotto A, Miller BA, Lewis DR, Eisner MP, Stinchcomb DG, Edwards BK (eds). SEER Cancer Statistics Review, 1975-2006, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2006/, based on November 2008 SEER data submission, posted to the SEER web site, 2009.
12. Vose JM. Mantle cell lymphoma: 2017 update on diagnosis, risk-stratification, and clinical management. *Am J Hematol*. 2017;92(8):806-813.
13. Stam K, Heisterkamp N, Grosveld G, et al. Evidence of a new chimeric bcr/c-abl mRNA in patients with chronic myelocytic leukemia and the Philadelphia chromosome. *N Engl J Med*. 1985;313(23):1429-1433.
14. Wang ZY, Chen Z. Acute promyelocytic leukemia: from highly fatal to highly curable. *Blood*. 2008;111(5):2505-2515.
15. Li S, Young KH, Medeiros LJ. Diffuse large B-cell lymphoma. *Pathology*. 2018;50(1):74-87.
16. Lossos IS, Gascoyne RD. Transformation of follicular lymphoma. *Best Pract Res Clin Haematol*. 2011;24(2):147-163.
17. Pasqualucci L, Dalla-Favera R. Genetics of diffuse large B-cell lymphoma. *Blood*. 2018;131(21):2307-2319.
18. Kuppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene*. 2001;20(40):5580-5594.
19. MacLennan IC. Germinal centers. *Annu Rev Immunol*. 1994;12:117-139.
20. Liu YJ, Zhang J, Lane PJ, Chan EY, MacLennan IC. Sites of specific B cell activation in primary and secondary responses to T cell-dependent and T cell-independent antigens. *Eur J Immunol*. 1991;21(12):2951-2962.

21. Schwickert TA, Victora GD, Fooksman DR, et al. A dynamic T cell-limited checkpoint regulates affinity-dependent B cell entry into the germinal center. *J Exp Med.* 2011;208(6):1243-1252.
22. Victora GD, Nussenzweig MC. Germinal centers. *Annu Rev Immunol.* 2012;30:429-457.
23. Nieuwenhuis P, Opstelten D. Functional anatomy of germinal centers. *Am J Anat.* 1984;170(3):421-435.
24. Schneider C, Pasqualucci L, Dalla-Favera R. Molecular pathogenesis of diffuse large B-cell lymphoma. *Semin Diagn Pathol.* 2011;28(2):167-177.
25. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem.* 2007;76:1-22.
26. Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. *Annu Rev Immunol.* 2008;26:261-292.
27. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell.* 2000;102(5):553-563.
28. Dorsett Y, Robbiani DF, Jankovic M, Reina-San-Martin B, Eisenreich TR, Nussenzweig MC. A role for AID in chromosome translocations between c-myc and the IgH variable region. *J Exp Med.* 2007;204(9):2225-2232.
29. Pasqualucci L, Migliazza A, Basso K, Houldsworth J, Chaganti RS, Dalla-Favera R. Mutations of the BCL6 proto-oncogene disrupt its negative autoregulation in diffuse large B-cell lymphoma. *Blood.* 2003;101(8):2914-2923.
30. Kitano M, Moriyama S, Ando Y, et al. Bcl6 protein expression shapes pre-germinal center B cell dynamics and follicular helper T cell heterogeneity. *Immunity.* 2011;34(6):961-972.
31. Basso K, Dalla-Favera R. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol Rev.* 2012;247(1):172-183.
32. Phan RT, Dalla-Favera R. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature.* 2004;432(7017):635-639.
33. Phan RT, Saito M, Basso K, Niu H, Dalla-Favera R. BCL6 interacts with the transcription factor Miz-1 to suppress the cyclin-dependent kinase inhibitor p21 and cell cycle arrest in germinal center B cells. *Nat Immunol.* 2005;6(10):1054-1060.
34. Ranuncolo SM, Polo JM, Dierov J, et al. Bcl-6 mediates the germinal center B cell phenotype and lymphomagenesis through transcriptional repression of the DNA-damage sensor ATR. *Nat Immunol.* 2007;8(7):705-714.
35. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol.* 2015;15(3):172-184.
36. Stevenson FK, Sahota SS, Ottensmeier CH, Zhu D, Forconi F, Hamblin TJ. The occurrence and significance of V gene mutations in B cell-derived human malignancy. *Adv Cancer Res.* 2001;83:81-116.
37. Lossos IS, Alizadeh AA, Eisen MB, et al. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc Natl Acad Sci U S A.* 2000;97(18):10209-10213.

38. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000;403(6769):503-511.
39. Lenz G, Staudt LM. Aggressive lymphomas. *N Engl J Med*. 2010;362(15):1417-1429.
40. Klein U, Casola S, Cattoretti G, et al. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat Immunol*. 2006;7(7):773-782.
41. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A*. 2003;100(17):9991-9996.
42. Savage KJ, Monti S, Kutok JL, et al. The molecular signature of mediastinal large B-cell lymphoma differs from that of other diffuse large B-cell lymphomas and shares features with classical Hodgkin lymphoma. *Blood*. 2003;102(12):3871-3879.
43. Rosenwald A, Wright G, Leroy K, et al. Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. *J Exp Med*. 2003;198(6):851-862.
44. Roschewski M, Staudt LM, Wilson WH. Diffuse large B-cell lymphoma-treatment approaches in the molecular era. *Nature reviews Clinical oncology*. 2014;11(1):12-23.
45. Pasqualucci L, Trifonov V, Fabbri G, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet*. 2011;43(9):830-837.
46. Lohr JG, Stojanov P, Lawrence MS, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A*. 2012;109(10):3879-3884.
47. Morin RD, Mungall K, Pleasance E, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood*. 2013;122(7):1256-1265.
48. Zhang J, Grubor V, Love CL, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A*. 2013;110(4):1398-1403.
49. Davis RE, Ngo VN, Lenz G, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature*. 2010;463(7277):88-92.
50. Ngo VN, Young RM, Schmitz R, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 2011;470(7332):115-119.
51. Lunning MA, Green MR. Mutation of chromatin modifiers; an emerging hallmark of germinal center B-cell lymphomas. *Blood Cancer J*. 2015;5:e361.
52. Fisher RI, Gaynor ER, Dahlberg S, et al. Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. *N Engl J Med*. 1993;328(14):1002-1006.
53. Coiffier B, Thieblemont C, Van Den Neste E, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood*. 2010;116(12):2040-2045.
54. Sehn LH, Donaldson J, Chhanabhai M, et al. Introduction of combined CHOP plus rituximab therapy dramatically improved outcome of diffuse large B-cell lymphoma in British Columbia. *J Clin Oncol*. 2005;23(22):5027-5033.

55. Habermann TM, Weller EA, Morrison VA, et al. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. *J Clin Oncol*. 2006;24(19):3121-3127.
56. Coiffier B, Lepage E, Briere J, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med*. 2002;346(4):235-242.
57. Cunningham D, Hawkes EA, Jack A, et al. Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles. *Lancet*. 2013;381(9880):1817-1826.
58. Purroy N, Bergua J, Gallur L, et al. Long-term follow-up of dose-adjusted EPOCH plus rituximab (DA-EPOCH-R) in untreated patients with poor prognosis large B-cell lymphoma. A phase II study conducted by the Spanish PETHEMA Group. *Br J Haematol*. 2015;169(2):188-198.
59. Lenz G, Wright G, Dave SS, et al. Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med*. 2008;359(22):2313-2323.
60. Thieblemont C, Briere J, Mounier N, et al. The germinal center/activated B-cell subclassification has a prognostic impact for response to salvage therapy in relapsed/refractory diffuse large B-cell lymphoma: a bio-CORAL study. *J Clin Oncol*. 2011;29(31):4079-4087.
61. Dabrowska-Iwanicka A, Walewski JA. Primary mediastinal large B-cell lymphoma. *Curr Hematol Malign Rep*. 2014;9(3):273-283.
62. Dunleavy K, Pittaluga S, Maeda LS, et al. Dose-adjusted EPOCH-rituximab therapy in primary mediastinal B-cell lymphoma. *N Engl J Med*. 2013;368(15):1408-1416.
63. Younes A, Ansell S, Fowler N, et al. The landscape of new drugs in lymphoma. *Nature reviews Clinical oncology*. 2017;14(6):335-346.
64. Aukema SM, Siebert R, Schuurin E, et al. Double-hit B-cell lymphomas. *Blood*. 2011;117(8):2319-2331.
65. Richmond TJ, Finch JT, Rushton B, Rhodes D, Klug A. Structure of the nucleosome core particle at 7 Å resolution. *Nature*. 1984;311(5986):532-537.
66. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res*. 2011;21(3):381-395.
67. Swygert SG, Peterson CL. Chromatin dynamics: interplay between remodeling enzymes and histone modifications. *Biochim Biophys Acta*. 2014;1839(8):728-736.
68. Clapier CR, Iwasa J, Cairns BR, Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol*. 2017;18(7):407-422.
69. Kadoch C, Hargreaves DC, Hodges C, et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet*. 2013;45(6):592-601.
70. Brownlee PM, Meisenberg C, Downs JA. The SWI/SNF chromatin remodelling complex: Its role in maintaining genome stability and preventing tumorigenesis. *DNA Repair (Amst)*. 2015;32:127-133.

71. Kakarougkas A, Ismail A, Chambers AL, et al. Requirement for PBAF in transcriptional repression and repair at DNA breaks in actively transcribed regions of chromatin. *Mol Cell*. 2014;55(5):723-732.
72. Ho L, Crabtree GR. Chromatin remodelling during development. *Nature*. 2010;463(7280):474-484.
73. de la Serna IL, Carlson KA, Imbalzano AN. Mammalian SWI/SNF complexes promote MyoD-mediated muscle differentiation. *Nat Genet*. 2001;27(2):187-190.
74. Centore RC, Sandoval GJ, Soares LMM, Kadoch C, Chan HM. Mammalian SWI/SNF Chromatin Remodeling Complexes: Emerging Mechanisms and Therapeutic Strategies. *Trends Genet*. 2020.
75. Hodges C, Kirkland JG, Crabtree GR. The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. *Cold Spring Harb Perspect Med*. 2016;6(8).
76. Alpsoy A, Dykhuizen EC. Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. *J Biol Chem*. 2018;293(11):3892-3903.
77. Mashtalir N, D'Avino AR, Michel BC, et al. Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell*. 2018;175(5):1272-1288 e1220.
78. Phelan ML, Sif S, Narlikar GJ, Kingston RE. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol Cell*. 1999;3(2):247-253.
79. Kadoch C, Crabtree GR. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv*. 2015;1(5):e1500447.
80. Hassan AH, Prochasson P, Neely KE, et al. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell*. 2002;111(3):369-379.
81. Hall JA, Georgel PT. CHD proteins: a diverse family with strong ties. *Biochem Cell Biol*. 2007;85(4):463-476.
82. Wang W, Xue Y, Zhou S, Kuo A, Cairns BR, Crabtree GR. Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev*. 1996;10(17):2117-2130.
83. Lessard J, Wu JI, Ranish JA, et al. An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron*. 2007;55(2):201-215.
84. Bogershausen N, Wollnik B. Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. *Front Mol Neurosci*. 2018;11:252.
85. Watanabe R, Ui A, Kanno S, et al. SWI/SNF factors required for cellular resistance to DNA damage include ARID1A and ARID1B and show interdependent protein stability. *Cancer Res*. 2014;74(9):2465-2475.
86. Chen Y, Zhang H, Xu Z, et al. A PARP1-BRG1-SIRT1 axis promotes HR repair by reducing nucleosome density at DNA damage sites. *Nucleic Acids Res*. 2019;47(16):8563-8580.
87. Dykhuizen EC, Hargreaves DC, Miller EL, et al. BAF complexes facilitate decatenation of DNA by topoisomerase IIalpha. *Nature*. 2013;497(7451):624-627.
88. Zhao B, Lin J, Rong L, et al. ARID1A promotes genomic stability through protecting telomere cohesion. *Nat Commun*. 2019;10(1):4067.

89. Kaeser MD, Aslanian A, Dong MQ, Yates JR, 3rd, Emerson BM. BRD7, a novel PBAF-specific SWI/SNF subunit, is required for target gene activation and repression in embryonic stem cells. *J Biol Chem*. 2008;283(47):32254-32263.
90. Mathur R, Alver BH, San Roman AK, et al. ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat Genet*. 2017;49(2):296-302.
91. Alver BH, Kim KH, Lu P, et al. The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nat Commun*. 2017;8:14648.
92. Wang X, Lee RS, Alver BH, et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat Genet*. 2017;49(2):289-295.
93. Nakayama RT, Pulice JL, Valencia AM, et al. SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. *Nat Genet*. 2017;49(11):1613-1623.
94. Biegel JA, Zhou JY, Rorke LB, Stenstrom C, Wainwright LM, Fogelgren B. Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors. *Cancer Res*. 1999;59(1):74-79.
95. Roberts CW, Galusha SA, McMenamin ME, Fletcher CD, Orkin SH. Haploinsufficiency of Snf5 (integrase interactor 1) predisposes to malignant rhabdoid tumors in mice. *Proc Natl Acad Sci U S A*. 2000;97(25):13796-13800.
96. Roberts CW, Leroux MM, Fleming MD, Orkin SH. Highly penetrant, rapid tumorigenesis through conditional inversion of the tumor suppressor gene Snf5. *Cancer Cell*. 2002;2(5):415-425.
97. Shain AH, Pollack JR. The spectrum of SWI/SNF mutations, ubiquitous in human cancers. *PLoS One*. 2013;8(1):e55119.
98. Wilson BG, Roberts CW. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer*. 2011;11(7):481-492.
99. Smith MJ, O'Sullivan J, Bhaskar SS, et al. Loss-of-function mutations in SMARCE1 cause an inherited disorder of multiple spinal meningiomas. *Nat Genet*. 2013;45(3):295-298.
100. Lu B, Shi H. An In-Depth Look at Small Cell Carcinoma of the Ovary, Hypercalcemic Type (SCCOHT): Clinical Implications from Recent Molecular Findings. *J Cancer*. 2019;10(1):223-237.
101. Varela I, Tarpey P, Raine K, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature*. 2011;469(7331):539-542.
102. Clark J, Rocques PJ, Crew AJ, et al. Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. *Nat Genet*. 1994;7(4):502-508.
103. Kadoch C, Crabtree GR. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. *Cell*. 2013;153(1):71-85.
104. Valencia AM, Kadoch C. Chromatin regulatory mechanisms and therapeutic opportunities in cancer. *Nat Cell Biol*. 2019;21(2):152-161.
105. Ogiwara H, Ui A, Otsuka A, et al. Histone acetylation by CBP and p300 at double-strand break sites facilitates SWI/SNF chromatin remodeling and the recruitment of non-homologous end joining factors. *Oncogene*. 2011;30(18):2135-2146.
106. Brownlee PM, Chambers AL, Cloney R, Bianchi A, Downs JA. BAF180 promotes cohesion and prevents genome instability and aneuploidy. *Cell Rep*. 2014;6(6):973-981.

107. Lee RS, Stewart C, Carter SL, et al. A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. *J Clin Invest*. 2012;122(8):2983-2988.
108. McBride MJ, Pulice JL, Beird HC, et al. The SS18-SSX Fusion Oncoprotein Hijacks BAF Complex Targeting and Function to Drive Synovial Sarcoma. *Cancer Cell*. 2018;33(6):1128-1141 e1127.
109. Hartwell LH, Szankasi P, Roberts CJ, Murray AW, Friend SH. Integrating genetic approaches into the discovery of anticancer drugs. *Science*. 1997;278(5340):1064-1068.
110. Kaelin WG, Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer*. 2005;5(9):689-698.
111. Schiaffino-Ortega S, Balinas C, Cuadros M, Medina PP. SWI/SNF proteins as targets in cancer therapy. *J Hematol Oncol*. 2014;7:81.
112. Oike T, Ogiwara H, Tominaga Y, et al. A synthetic lethality-based strategy to treat cancers harboring a genetic deficiency in the chromatin remodeling factor BRG1. *Cancer Res*. 2013;73(17):5508-5518.
113. Wilson BG, Helming KC, Wang X, et al. Residual complexes containing SMARCA2 (BRM) underlie the oncogenic drive of SMARCA4 (BRG1) mutation. *Mol Cell Biol*. 2014;34(6):1136-1144.
114. Hoffman GR, Rahal R, Buxton F, et al. Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. *Proc Natl Acad Sci U S A*. 2014;111(8):3128-3133.
115. Helming KC, Wang X, Wilson BG, et al. ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat Med*. 2014;20(3):251-254.
116. Michel BC, D'Avino AR, Cassel SH, et al. A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. *Nat Cell Biol*. 2018;20(12):1410-1420.
117. Wang X, Wang S, Troisi EC, et al. BRD9 defines a SWI/SNF sub-complex and constitutes a specific vulnerability in malignant rhabdoid tumors. *Nat Commun*. 2019;10(1):1881.
118. Vangamudi B, Paul TA, Shah PK, et al. The SMARCA2/4 ATPase Domain Surpasses the Bromodomain as a Drug Target in SWI/SNF-Mutant Cancers: Insights from cDNA Rescue and PFI-3 Inhibitor Studies. *Cancer Res*. 2015;75(18):3865-3878.
119. Papillon JPN, Nakajima K, Adair CD, et al. Discovery of Orally Active Inhibitors of Brahma Homolog (BRM)/SMARCA2 ATPase Activity for the Treatment of Brahma Related Gene 1 (BRG1)/SMARCA4-Mutant Cancers. *J Med Chem*. 2018;61(22):10155-10172.
120. Bondeson DP, Mares A, Smith IE, et al. Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat Chem Biol*. 2015;11(8):611-617.
121. Buscarlet M, Krasteva V, Ho L, et al. Essential role of BRG, the ATPase subunit of BAF chromatin remodeling complexes, in leukemia maintenance. *Blood*. 2014;123(11):1720-1728.
122. Shi J, Whyte WA, Zepeda-Mendoza CJ, et al. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev*. 2013;27(24):2648-2662.
123. Hohmann AF, Martin LJ, Minder JL, et al. Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition. *Nat Chem Biol*. 2016;12(9):672-679.

124. Zani VJ, Asou N, Jadayel D, et al. Molecular cloning of complex chromosomal translocation t(8;14;12)(q24.1;q32.3;q24.1) in a Burkitt lymphoma cell line defines a new gene (BCL7A) with homology to caldesmon. *Blood*. 1996;87(8):3124-3134.
125. Nacheva E, Fischer P, Karpas A, et al. Complex translocation t(8;12;14) in a cell line derived from a child with nonendemic Burkitt-type acute lymphoblastic leukemia. *Cancer Genet Cytogenet*. 1987;28(1):145-153.
126. Willis TG, Dyer MJ. The role of immunoglobulin translocations in the pathogenesis of B-cell malignancies. *Blood*. 2000;96(3):808-822.
127. Jadayel DM, Osborne LR, Coignet LJ, et al. The BCL7 gene family: deletion of BCL7B in Williams syndrome. *Gene*. 1998;224(1-2):35-44.
128. van Doorn R, Zoutman WH, Dijkman R, et al. Epigenetic profiling of cutaneous T-cell lymphoma: promoter hypermethylation of multiple tumor suppressor genes including BCL7a, PTPRG, and p73. *J Clin Oncol*. 2005;23(17):3886-3896.
129. Tracey L, Villuendas R, Dotor AM, et al. Mycosis fungoides shows concurrent deregulation of multiple genes involved in the TNF signaling pathway: an expression profile study. *Blood*. 2003;102(3):1042-1050.
130. Carbone A, Bernardini L, Valenzano F, et al. Array-based comparative genomic hybridization in early-stage mycosis fungoides: recurrent deletion of tumor suppressor genes BCL7A, SMAC/DIABLO, and RHOA. *Genes Chromosomes Cancer*. 2008;47(12):1067-1075.
131. Martinez-Delgado B, Melendez B, Cuadros M, et al. Expression profiling of T-cell lymphomas differentiates peripheral and lymphoblastic lymphomas and defines survival related genes. *Clin Cancer Res*. 2004;10(15):4971-4982.
132. Ho L, Ronan JL, Wu J, et al. An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A*. 2009;106(13):5181-5186.
133. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011;476(7360):298-303.
134. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature*. 2011;471(7339):467-472.
135. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014;505(7484):495-501.
136. Reddy A, Zhang J, Davis NS, et al. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell*. 2017;171(2):481-494 e415.
137. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N Engl J Med*. 2018;378(15):1396-1407.
138. Krysiak K, Gomez F, White BS, et al. Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. *Blood*. 2017;129(4):473-483.
139. Pasqualucci L, Guglielmino R, Houldsworth J, et al. Expression of the AID protein in normal and neoplastic B cells. *Blood*. 2004;104(10):3318-3325.
140. Shaffer AL, Yu X, He Y, Boldrick J, Chan EP, Staudt LM. BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity*. 2000;13(2):199-212.

141. Ramos-Medina R, Montes-Moreno S, Maestre L, et al. BCL7A protein expression in normal and malignant lymphoid tissues. *Br J Haematol*. 2013;160(1):106-109.
142. Heng TS, Painter MW, Immunological Genome Project C. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9(10):1091-1094.
143. Kato L, Begum NA, Burroughs AM, et al. Nonimmunoglobulin target loci of activation-induced cytidine deaminase (AID) share unique features with immunoglobulin genes. *Proc Natl Acad Sci U S A*. 2012;109(7):2479-2484.
144. Khodabakhshi AH, Morin RD, Fejes AP, et al. Recurrent targets of aberrant somatic hypermutation in lymphoma. *Oncotarget*. 2012;3(11):1308-1319.
145. Storb U. Why does somatic hypermutation by AID require transcription of its target genes? *Adv Immunol*. 2014;122:253-277.
146. Vicente-Duenas C, Fontan L, Gonzalez-Herrero I, et al. Expression of MALT1 oncogene in hematopoietic stem/progenitor cells recapitulates the pathogenesis of human lymphoma in mice. *Proc Natl Acad Sci U S A*. 2012;109(26):10534-10539.
147. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
148. Alvarez-Prado AF, Perez-Duran P, Perez-Garcia A, et al. A broad atlas of somatic hypermutation allows prediction of activation-induced deaminase targets. *J Exp Med*. 2018.
149. Naldini L, Blomer U, Gallay P, et al. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*. 1996;272(5259):263-267.
150. Benabdellah K, Gutierrez-Guerrero A, Cobo M, Munoz P, Martin F. A chimeric HS4-SAR insulator (IS2) that prevents silencing and enhances expression of lentiviral vectors in pluripotent stem cells. *PLoS One*. 2014;9(1):e84268.
151. Eekels JJ, Pasternak AO, Schut AM, Geerts D, Jeeninga RE, Berkhout B. A competitive cell growth assay for the detection of subtle effects of gene transduction on cell proliferation. *Gene Ther*. 2012;19(11):1058-1064.
152. Pasqualucci L, Khiabani H, Fangazio M, et al. Genetics of follicular lymphoma transformation. *Cell Rep*. 2014;6(1):130-140.
153. Carmel I, Tal S, Vig I, Ast G. Comparative analysis detects dependencies among the 5' splice-site positions. *RNA*. 2004;10(5):828-840.
154. Grande BM, Gerhard DS, Jiang A, et al. Genome-wide discovery of somatic coding and noncoding mutations in pediatric endemic and sporadic Burkitt lymphoma. *Blood*. 2019;133(12):1313-1324.
155. Pasqualucci L, Neumeister P, Goossens T, et al. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature*. 2001;412(6844):341-346.
156. Perez-Duran P, Belver L, de Yébenes VG, Delgado P, Pisano DG, Ramiro AR. UNG shapes the specificity of AID-induced somatic hypermutation. *J Exp Med*. 2012;209(7):1379-1389.
157. Rada C, Di Noia JM, Neuberger MS. Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol Cell*. 2004;16(2):163-171.

158. Methot SP, Di Noia JM. Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. *Adv Immunol.* 2017;133:37-87.
159. Rogozin IB, Kolchanov NA. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim Biophys Acta.* 1992;1171(1):11-18.
160. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun.* 2019;10(1):1523.
161. Velichutina I, Shaknovich R, Geng H, et al. EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis. *Blood.* 2010;116(24):5247-5255.
162. Chae YS, Kim H, Kim D, Lee H, Lee HO. Cell density-dependent acetylation of DeltaNp63alpha is associated with p53-dependent cell cycle arrest. *FEBS Lett.* 2012;586(8):1128-1134.
163. Tan SH, Yam AW, Lawton LN, et al. TRIB2 reinforces the oncogenic transcriptional program controlled by the TAL1 complex in T-cell acute lymphoblastic leukemia. *Leukemia.* 2016;30(4):959-962.
164. Hou Z, Guo K, Sun X, et al. TRIB2 functions as novel oncogene in colorectal cancer by blocking cellular senescence through AP4/p21 signaling. *Mol Cancer.* 2018;17(1):172.
165. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 2016;44(W1):W90-97.
166. Wilson BG, Wang X, Shen X, et al. Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. *Cancer Cell.* 2010;18(4):316-328.
167. Bitler BG, Aird KM, Garipov A, et al. Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. *Nat Med.* 2015;21(3):231-238.
168. Knutson SK, Warholic NM, Wigle TJ, et al. Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. *Proc Natl Acad Sci U S A.* 2013;110(19):7922-7927.
169. Morin RD, Johnson NA, Severson TM, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet.* 2010;42(2):181-185.
170. Beguelin W, Popovic R, Teater M, et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell.* 2013;23(5):677-692.
171. Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet.* 2007;8(10):749-761.
172. Cooper TA, Wan L, Dreyfuss G. RNA and disease. *Cell.* 2009;136(4):777-793.
173. Singh RK, Cooper TA. Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med.* 2012;18(8):472-482.
174. Krawczak M, Thomas NS, Hundrieser B, et al. Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Hum Mutat.* 2007;28(2):150-158.
175. Gibbons, R., Higgs, D. R., Olivieri, N. F. & Wood, W. G. in *The Thalassemia Syndromes* 4th edn Ch. 7 (eds Weatherall, D. J. & Clegg, J. B.) 287–356 (Blackwell Science, Oxford, 2001).

176. Srebrow A, Kornblihtt AR. The connection between splicing and cancer. *J Cell Sci.* 2006;119(Pt 13):2635-2641.
177. Weinhold N, Ashby C, Rasche L, et al. Clonal selection and double-hit events involving tumor suppressor genes underlie relapse in myeloma. *Blood.* 2016;128(13):1735-1744.
178. Panea RI, Love CL, Shingleton JR, et al. The whole-genome landscape of Burkitt lymphoma subtypes. *Blood.* 2019;134(19):1598-1607.
179. Liu M, Duke JL, Richter DJ, et al. Two levels of protection for the B cell genome during somatic hypermutation. *Nature.* 2008;451(7180):841-845.
180. Pasqualucci L, Migliazza A, Fracchiolla N, et al. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci U S A.* 1998;95(20):11816-11821.
181. Shen HM, Peters A, Baron B, Zhu X, Storb U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science.* 1998;280(5370):1750-1752.
182. Zhao K, Wang W, Rando OJ, et al. Rapid and phosphoinositol-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. *Cell.* 1998;95(5):625-636.
183. Mashtalir N, Suzuki H, Farrell DP, et al. A Structural Model of the Endogenous Human BAF Complex Informs Disease Mechanisms. *Cell.* 2020.
184. He S, Wu Z, Tian Y, et al. Structure of nucleosome-bound human BAF complex. *Science.* 2020;367(6480):875-881.
185. Bossen C, Murre CS, Chang AN, Mansson R, Rodewald HR, Murre C. The chromatin remodeler Brg1 activates enhancer repertoires to establish B cell identity and modulate cell growth. *Nat Immunol.* 2015;16(7):775-784.
186. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science.* 2017;357(6348).
187. Allshire RC, Madhani HD. Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol.* 2018;19(4):229-244.
188. Liu Y, Gonzalez Y, Amengual JE. Chromatin-Remodeled State in Lymphoma. *Curr Hematol Malig Rep.* 2019;14(5):439-450.
189. Holley DW, Groh BS, Wozniak G, et al. The BRG1 chromatin remodeler regulates widespread changes in gene expression and cell proliferation during B cell activation. *J Cell Physiol.* 2014;229(1):44-52.
190. Kadoch C, Williams RT, Calarco JP, et al. Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states. *Nat Genet.* 2017;49(2):213-222.
191. Stanton BZ, Hodges C, Calarco JP, et al. Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. *Nat Genet.* 2017;49(2):282-288.
192. Beguelin W, Teater M, Gearhart MD, et al. EZH2 and BCL6 Cooperate to Assemble CBX8-BCOR Complex to Repress Bivalent Promoters, Mediate Germinal Center Formation and Lymphomagenesis. *Cancer Cell.* 2016;30(2):197-213.

193. Beguelin W, Rivas MA, Calvo Fernandez MT, et al. EZH2 enables germinal centre formation through epigenetic silencing of CDKN1A and an Rb-E2F1 feedback loop. *Nat Commun.* 2017;8(1):877.
194. Choi J, Jeon S, Choi S, Park K, Seong RH. The SWI/SNF chromatin remodeling complex regulates germinal center formation by repressing Blimp-1 expression. *Proc Natl Acad Sci U S A.* 2015;112(7):E718-727.
195. Sood S, Weber CM, Hodges HC, Krokhotin A, Shalizi A, Crabtree GR. CHD8 dosage regulates transcription in pluripotency and early murine neural differentiation. *Proc Natl Acad Sci U S A.* 2020;117(36):22331-22340.
196. Parekh S, Polo JM, Shaknovich R, et al. BCL6 programs lymphoma cells for survival and differentiation through distinct biochemical mechanisms. *Blood.* 2007;110(6):2067-2074.
197. Reynolds N, Salmon-Divon M, Dvinge H, et al. NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *EMBO J.* 2012;31(3):593-605.
198. Cattoretti G, Pasqualucci L, Ballon G, et al. Deregulated BCL6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice. *Cancer Cell.* 2005;7(5):445-455.
199. Berg T, Thoene S, Yap D, et al. A transgenic mouse model demonstrating the oncogenic role of mutations in the polycomb-group gene EZH2 in lymphomagenesis. *Blood.* 2014;123(25):3914-3924.
200. Kim KH, Kim W, Howard TP, et al. SWI/SNF-mutant cancers depend on catalytic and non-catalytic activity of EZH2. *Nat Med.* 2015;21(12):1491-1496.
201. Knutson SK, Warholic NM, Johnston LD, et al. Synergistic Anti-Tumor Activity of EZH2 Inhibitors and Glucocorticoid Receptor Agonists in Models of Germinal Center Non-Hodgkin Lymphomas. *PLoS One.* 2014;9(12):e111840.
202. Italiano A, Soria JC, Toulmonde M, et al. Tazemetostat, an EZH2 inhibitor, in relapsed or refractory B-cell non-Hodgkin lymphoma and advanced solid tumours: a first-in-human, open-label, phase 1 study. *Lancet Oncol.* 2018;19(5):649-659.
203. Mittal P, Roberts CWM. The SWI/SNF complex in cancer - biology, biomarkers and therapy. *Nature reviews Clinical oncology.* 2020.

9. DERECHOS DE AUTOR Y PERMISOS

Figura 1

License Number 4936480247782

License date Oct 26, 2020

📄 Licensed Content

Licensed Content Publisher Elsevier
 Licensed Content Publication Seminars in Diagnostic Pathology
 Licensed Content Title Molecular pathogenesis of diffuse large B-cell lymphoma
 Licensed Content Author Christof Schneider, Laura Pasqualucci, Riccardo Dalla-Favera
 Licensed Content Date May 1, 2011
 Licensed Content Volume 28
 Licensed Content Issue 2
 Licensed Content Pages 11
 Journal Type HS

📄 Order Details

Type of Use reuse in a thesis/dissertation
 Portion figures/tables/illustrations
 Number of figures/tables/illustrations 1
 Format both print and electronic
 Are you the author of this Elsevier article? No
 Will you be translating? No

Figura 2

License Number 4936480813365

License date Oct 26, 2020

📄 Licensed Content

Licensed Content Publisher Springer Nature
 Licensed Content Publication Nature Reviews Immunology
 Licensed Content Title Germinal centres and B cell lymphomagenesis
 Licensed Content Author Katia Basso et al
 Licensed Content Date Feb 25, 2015

📄 Order Details

Type of Use Thesis/Dissertation
 Requestor type academic/university or research institute
 Format print and electronic
 Portion figures/tables/illustrations
 Number of figures/tables/illustrations 2
 High-res required no
 Will you be translating? no
 Circulation/distribution 1 - 29
 Author of this Springer Nature content no

Figura 3

License Number 4936481245977

License date Oct 26, 2020

📄 Licensed Content

Licensed Content Publisher Springer Nature
 Licensed Content Publication Nature Reviews Molecular Cell Biology
 Licensed Content Title Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes
 Licensed Content Author Cedric R. Clapier et al
 Licensed Content Date May 17, 2017

📄 Order Details

Type of Use Thesis/Dissertation
 Requestor type academic/university or research institute
 Format print and electronic
 Portion figures/tables/illustrations
 Number of figures/tables/illustrations 1
 High-res required no
 Will you be translating? no
 Circulation/distribution 1 - 29
 Author of this Springer Nature content no

Figura 4

License Number 4936490212452

License date Oct 26, 2020

 Licensed Content

Licensed Content Publisher	Elsevier
Licensed Content Publication	Cell
Licensed Content Title	Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes
Licensed Content Author	Nazar Mashtali, Andrew R. D'Avino, Brittany C. Michel, Jie Luo, Joshua Pan, Jordan E. Otto, Hayley J. Zullo, Zachary M. McKenzie, Rachel L. Kubiak, Roodolph St. Pierre, Alfredo M. Valencia, Steven J. Poynter, Seth H. Cassel, Jeffrey A. Ranish, Cigall Kadoch
Licensed Content Date	Nov 15, 2018
Licensed Content Volume	175
Licensed Content Issue	5
Licensed Content Pages	37
Journal Type	S&T

 Order Details

Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No

Figura 6

License Number 4936491431258

License date Oct 26, 2020

 Licensed Content

Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Cell Biology
Licensed Content Title	Chromatin regulatory mechanisms and therapeutic opportunities in cancer
Licensed Content Author	Alfredo M. Valencia et al
Licensed Content Date	Jan 2, 2019

 Order Details

Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no

10. APÉNDICES

APÉNDICE I

La tabla que se muestra en el apéndice I busca servir de referencia para reconocer cada una de las subunidades de los complejo SWI/SNF en humanos a través de las diferentes nomenclaturas que tienen. Primero, se introdujo la nomenclatura con el acrónimo BAF (*BRG1/BRM-associated factor*) seguido de un número que indica el tamaño de las proteínas (en KDa) codificadas por los distintos genes. Posteriormente, la nomenclatura oficial HUGO (**Human Genome Organization**) se realizó a partir del acrónimo SMARC (**S**WI/SNF-related, **M**atrix-associated, **A**ctin-dependent **R**egulator of **C**hromatin). Aun así, las evidencias detrás de la relación del complejo SWI/SNF con la matriz extracelular tenían ciertas limitaciones y muchas de las subunidades descritas más recientemente han adoptado nombres propios.

En el Apéndice I se muestran también los distintos complejos BAF que forman cada una de las subunidades: canonical BAF (cBAF), polybromo-associated BAF (PFAF) o non-canonical BAF (ncBAF). Por último, en la última columna se muestra el módulo estructural en el que se ha caracterizado a cada una de las subunidades según el estudio llevado a cabo por Mashtalir, N. et al. (2018)⁷⁷.

La tabla se ha adaptado de una de las últimas revisiones publicadas del complejo SWI/SNF⁷⁴. Las subunidades BCL11A y BCL11B no se han tenido en cuenta como subunidades del complejo ya que no se han encontrado asociadas al complejo en diversos estudios proteómicos, como es el caso de esta tesis doctoral. Aun así, otros autores siguen considerando a estas subunidades como integrantes del complejo SWI/SNF²⁰³.

APÉNDICE II

La tabla que se muestra en el apéndice II muestra todas y cada una de las líneas celulares derivadas de DLBCL que han sido utilizadas en el estudio de secuenciación de *BCL7A*. En la tabla se muestran los repositorios de líneas celulares del cual se han obtenido cada una de las líneas celulares, así como los números de identificación. Algunas de las líneas celulares no están disponibles en los repositorios y se indica la referencia bibliográfica que indica el origen y los investigadores que han cedido esas líneas celulares.

APÉNDICE III

La tabla que se muestra en el apéndice III resume las variaciones genéticas encontradas para el gen *BCL7A* tanto en la colección de 41 líneas celulares como en la cohorte de pacientes analizadas en este estudio. Se indica la posición de las mutaciones tanto a nivel de ADN, como a nivel de proteína. Se indica también, si se conoce, el subtipo de DLBCL en función de la célula de origen.

APÉNDICE IV

En el apéndice IV se muestra una tabla con algunas de las proteínas que fueron detectadas mediante los análisis de LC/MS-MS a partir de los inmunocomplejos de BCL7A en la línea celular HBL1 (Ver sección materiales y métodos). En la tabla se muestra tanto el nombre de la proteína como el ID con el cual se encuentran anotadas en las bases de datos.

Las proteínas que se muestran en la tabla se han ordenado en función a los valores decrecientes del *fold change* (FC) para la condición $\Delta 27$ -BCL7A vs IgG siempre y cuando los valores fuesen $\log_2 FC > 1,5$. Además, se rescataron aquellas proteínas que tenían también valores de $\log_2 FC > 1,5$ para la condición BCL7A salvaje (WT) vs IgG. Las celdas de las dos condiciones mencionadas que presentan valores $\log_2 FC > |1|$ y un p-valor $< 0,05$ aparecen coloreadas. Se muestra también una columna en la que aparecen los valores de FC calculados por el software *Limma* para la comparación $\Delta 27$ -BCL7A vs WT.

Usando las 134 proteínas que se muestran en la tabla, se realizó un análisis de enriquecimiento mediante la herramienta STRING. En una de las columnas de la tabla se muestran algunos de los componentes celulares obtenidos en el análisis de enriquecimiento por ontología génica. Además de haber obtenido un enriquecimiento esperado para las subunidades de los complejos SWI/SNF, aparece representado otro complejo remodelador de la cromatina como es el complejo NuRD y un complejo relacionado con la polimerización de actina, el complejo Arp2/3. Además, en la última columna se muestra una flecha que indica si las proteínas que pertenecen a los componentes celulares mencionados se detectan en mayor (\uparrow) o menor (\downarrow) cantidad al comparar $\Delta 27$ -BCL7A vs WT.

Por último, algunas celdas de proteínas han sido marcadas en amarillo. Estas proteínas parecen unirse de manera similar (=) tanto a la forma mutante $\Delta 27$ -BCL7A como a la forma WT. Todas ellas tienen valores de $\log_2 FC > 1,5$ tanto para la comparación $\Delta 27$ -BCL7A vs IgG como para la comparación WT vs IgG. Además, tienen un valor de $\log_2 FC < |1|$ al comparar $\Delta 27$ -BCL7A vs WT.

APÉNDICE V

Portada del artículo publicado en la revista *Leukemia* donde se incluyen gran parte de los resultados generados durante la realización de esta tesis doctoral. En el apéndice V solamente se muestra una imagen con el resumen del artículo junto con los nombres completos y afiliaciones de todos los autores que han colaborado en su realización.

APÉNDICE I. Subunidades de los complejos SWI/SNF en mamíferos

Gen	Alias	Complejos mSWI/SNF	Módulo
ACTB	β -actina	cBAF / PBAF/ncBAF	ATPasa
ACTL6A	BAF53A	cBAF / PBAF/ncBAF	ATPasa
ACTL6B	BAF53B	cBAF / PBAF/ncBAF	ATPasa
BCL7A		cBAF / PBAF/ncBAF	ATPasa
BCL7B		cBAF / PBAF/ncBAF	ATPasa
BCL7C		cBAF / PBAF/ncBAF	ATPasa
SMARCA2	BRM	cBAF / PBAF/ncBAF	ATPasa
SMARCA4	BRG1	cBAF / PBAF/ncBAF	ATPasa
SMARCC1	BAF155	cBAF / PBAF/ncBAF	Núcleo BAF
SMARCC2	BAF170	cBAF / PBAF/ncBAF	Núcleo BAF
SMARCD1	BAF60A	cBAF / PBAF/ncBAF	Núcleo BAF
SMARCD2	BAF60B	cBAF / PBAF/ncBAF	Núcleo BAF
SMARCD3	BAF60C	cBAF / PBAF/ncBAF	Núcleo BAF
ARID1A	BAF250A	cBAF	Específica cBAF
ARID1B	FAB250B	cBAF	Específica cBAF
ARID2	BAF200	PBAF	Específica PBAF
BICRA	GLTSCR1	ncBAF	Específica ncBAF
BICRAL	GLTSCRL1	ncBAF	Específica ncBAF
BRD7		PBAF	Específica cBAF
BRD9		ncBAF	Específica ncBAF
DPF1	BAF45B	cBAF	Específica cBAF
DPF2	BAF45D	cBAF	Específica cBAF
DPF3	BAF45C	cBAF	Específica cBAF
PHF10	BAF45A	PBAF	Específica PBAF
PBRM1	BAF180	PBAF	Específica PBAF
SMARCB1	BAF47 / INI1	cBAF / PBAF	Núcleo BAF
SMARCE1	BAF57	cBAF / PBAF	Núcleo BAF
SS18	SSXT / SYT	cBAF / ncBAF	ATPasa
SS18L1	CREST	cBAF / ncBAF	ATPasa

APÉNDICE II. Variaciones génicas en BCL7A en nuestra cohorte de pacientes y líneas celulares

Muestra	Linfoma subtipo	Localización	Mutaciones		Impacto funcional
			Nucleótido	Cambio AA	
Mutaciones de <i>splicing</i>					
OCI-LY1	GCB-DLBCL	Intrón 1 (<i>splicing</i>)	92+1G>T	S5_W31del	Δ27-BCL7A
Karpas-1106	PMBCL	Intrón 1 (<i>splicing</i>)	92+1G>C	S5_W31del	Δ27-BCL7A
VAL	GCB-DLBCL	Intrón 1 (<i>splicing</i>)	92+1G>C	S5_W31del	Δ27-BCL7A
F13 (Paciente)	FL	Intrón 1 (<i>splicing</i>)	92+1G>C	S5_W31del	Δ27-BCL7A
Mutaciones con cambio de sentido					
OCI-LY7	GCB-DLBCL	Exón 1	85C>T	R29C	Cambio de aminoácido
DB	GCB-DLBCL	Exón 1	33G>C	R11S	Cambio de aminoácido
VAL	GCB-DLBCL	Exón 1	32G>C	R11T	Cambio de aminoácido
ROS-50	GCB-DLBCL	Exón 1	[23C>A] + [86G>T]	A8D + R29L	Cambio de aminoácido
D32 (Paciente)	GCB-DLBCL	Exón 1	17T>C	V6A	Cambio de aminoácido
D29 (Paciente)	No clasificado	Exón 1	32G>T	R11M	Cambio de aminoácido
Mutaciones con cambio de marco de lectura					
OCI-LY1	GCB-DLBCL	Exón 1	2T>G	M1R	Pérdida codón de inicio
Translocaciones cromosómicas y deleciones					
Karpas-1106	PMBCL	Exones 2 y 3	Duplicación en tándem	--	Codón de terminación prematuro
HBL1	ABC-DLBCL	Exón 1	Deleción homocigota	--	No expresa BCL7A

APÉNDICE III. Líneas celulares de DLBCL

Línea celular	ATCC	DSMZ	Origen / Descripción
DB	CRL-2289		
DOHH-2		ACC 47	
Farage	CRL-2630		
Granta-452		ACC 713	
HBL1			Nozawa, Y. et al., 1988
HLY1			Brown, P. J. et al., 2008
HT		ACC 567	
Karpas-231			Nacheva, E. et al., 1993
Karpas-422			Dyer, M. J. et al., 1990
Karpas-1106			Nacheva, E. et al., 1994
MD901			Miki, T. et al., 1994
MEDB1			Bentz, M. et al., 2001
NU-DHL-1		ACC 583	
NU-DUL-1	CRL-2969		
OCI-LY1		ACC 722	
OCI-LY3		ACC 761	
OCI-LY4			Mehra, S. et al., 200
OCI-LY7		ACC 688	
OCI-LY8			Tweeddale, M. E. et al., 1987
OCI-LY10			Tweeddale, M. E. et al., 1987
OCI-LY19		ACC 528	
OZ			
Pfeiffer	CRL-2632		
RC-K8		ACC 561	
PR1			Mestre-Escorihuela, C. et al., 2007
RIVA		ACC 585	
RL		ACC 613	
ROS-50		ACC 557	
SC-1		ACC 558	
SU-DHL-4		ACC 495	
SU-DHL-5		ACC 571	
SU-DHL-6	CRL-2959		
SU-DHL-8	CRL-2961		
SU-DHL-10		ACC 576	
SU-DHL-16	CRL-2964		
Toledo	CRL-2631		
U-2932		ACC 633	
VAL		ACC 586	
WSU-DLCL2		ACC 575	
WSU-NHL		ACC 58	
380		ACC 39	

Bentz M, Barth TF, Bruderlein S, Bock D, Schwerer MJ, Baudis M, et al. Gain of chromosome arm 9p is characteristic of primary mediastinal B-cell lymphoma (MBL): comprehensive molecular cytogenetic analysis and presentation of a novel MBL cell line. *Genes Chromosomes Cancer*. **2001**;30(4):393-401.

Brown PJ, Ashe SL, Leich E, Burek C, Barrans S, Fenton JA, et al. Potentially oncogenic B-cell activation-induced smaller isoforms of FOXP1 are highly expressed in the activated B cell-like subtype of DLBCL. *Blood*. **2008**;111(5):2816-24.

Dyer MJ, Fischer P, Nacheva E, Labastide W, Karpas A. A new human B-cell non-Hodgkin's lymphoma cell line (Karpas 422) exhibiting both t(14;18) and t(4;11) chromosomal translocations. *Blood*. **1990**;75(3):709-14.

Mehra S, Messner H, Minden M, Chaganti RS. Molecular cytogenetic characterization of non-Hodgkin lymphoma cell lines. *Genes Chromosomes Cancer*. **2002**;33(3):225-34.

Mestre-Escorihuela C, Rubio-Moscardo F, Richter JA, Siebert R, Climent J, Fresquet V, et al. Homozygous deletions localize novel tumor suppressor genes in B-cell lymphomas. *Blood*. **2007**;109(1):271-80.

Miki T, Kawamata N, Arai A, Ohashi K, Nakamura Y, Kato A, et al. Molecular cloning of the breakpoint for 3q27 translocation in B-cell lymphomas and leukemias. *Blood*. **1994**;83(1):217-22.

Nacheva E, Dyer MJ, Fischer P, Stranks G, Heward JM, Marcus RE, et al. C-MYC translocations in de novo B-cell lineage acute leukemias with t(14;18)(cell lines Karpas 231 and 353). *Blood*. **1993**;82(1):231-40.

Nacheva E, Dyer MJ, Metivier C, Jadayel D, Stranks G, Morilla R, et al. B-cell non-Hodgkin's lymphoma cell line (Karpas 1106) with complex translocation involving 18q21.3 but lacking BCL2 rearrangement and expression. *Blood*. **1994**;84(10):3422-8.

Nozawa Y, Abe M, Wakasa H, Ohno H, Fukuhara S, Kinoshita T, et al. Establishment and characterization of an Epstein-Barr virus negative B-cell lymphoma cell line and successful heterotransplantation. *The Tohoku journal of experimental medicine*. **1988**;156(4):319-30.

Tweeddale ME, Lim B, Jamal N, Robinson J, Zalcborg J, Lockwood G, et al. The presence of clonogenic cells in high-grade malignant lymphoma: a prognostic factor. *Blood*. **1987**;69(5):1307-14.

APÉNDICE IV. Resultados obtenidos por espectrometría de masas (LC/MS-MS)

Proteína	ID Proteína	$\Delta 27$ vs IgG	WT vs IgG	$\Delta 27$ vs WT	Componente celular (Gene Ontology)	Cambio MUT vs WT
SMARCC2	Q8TAQ2	3.13	14.53	-12.66	SWI/SNF	↓
SMARCA2	P51531	1.51	14.04	-13.03	SWI/SNF	↓
ACTL6A	O96019	2.25	13.81	-11.98	SWI/SNF	↓
DPF2	Q92785	-2.32	13.80	-12.07	SWI/SNF	↓
SMARCC1	Q92922	0.92	12.83	-11.51	SWI/SNF	↓
SMARCA4	P51532	-2.60	11.74	-11.80	SWI/SNF	↓
PBRM1	Q86U86	-2.35	11.47	-11.38	SWI/SNF	↓
PAWR	Q96I20	10.64	11.42	0.95		=
SMARCB1	Q12824	0.74	11.02	-10.05	SWI/SNF	↓
ARID1A	O14497	2.60	10.84	-9.20	SWI/SNF	↓
SMARCE1	Q969G3	1.78	10.21	-7.89	SWI/SNF	↓
BCL7A	Q4VC05	13.50	10.05	2.91	SWI/SNF	
SMARCD2	Q92925	-0.86	9.61	-10.61	SWI/SNF	↓
ARID1B	Q8NFD5	0.61	9.25	-8.53	SWI/SNF	↓
SMARCD1	Q96GM5	-1.23	8.47	-10.03	SWI/SNF	↓
ARID2	Q68CP9	0.06	8.41	-8.23	SWI/SNF	↓
PHF10	Q8WUB8	1.01	8.18	-7.05	SWI/SNF	↓
SS18	Q15532	0.06	7.95	-7.15	SWI/SNF	↓
BRD9	Q9H8M2	-0.52	7.54	-7.95	SWI/SNF	↓
BRD7	Q9NPI1	1.58	7.36	-5.99	SWI/SNF	↓
GLTSCR1L	Q6AI39	0.06	7.29	-7.11	SWI/SNF	↓
ACAP2	Q15057	6.17	6.60	-0.31		=
GLTSCR1	Q9NZM4	0.06	6.33	-6.15	SWI/SNF	↓
SS18L1	O75177	0.06	5.06	-4.88	SWI/SNF	↓
ACTA1	P68133	-0.70	4.76	-4.49	Actina	↓
ENO3	P13929	0.06	4.18	-4.00		
DPF3	Q92784	0.06	4.16	-3.98		
SNX18	Q96RF0	2.50	3.85	-1.24		
NPM3	O75607	0.98	3.79	-1.01		
ATP5F1	P24539	3.56	3.77	-0.10		=
CHD4	Q14839	0.85	3.74	-2.13	NuRD complex	↓
HDAC1	Q13547	2.32	3.71	-1.27	NuRD complex	↓
CSNK2A1	P68400	0.45	3.53	-2.86	NuRD complex	↓
KPNA3	O00505	0.28	3.26	-3.24		
CSNK2A2	P19784	0.21	3.22	-2.96		
RBBP4	Q09028	0.97	3.20	-1.83	NuRD complex	↓
C4A	POCOL4	3.74	3.16	0.89		=
KPNA2	P52292	0.87	3.08	-2.11		
TMEM159	Q96B96	2.82	3.01	-0.35		=
ZBTB37	Q5TC79	3.29	2.91	0.50		=
CEP97	Q8IW35	3.01	2.85	0.28		=
MYL6B	P14649	0.06	2.70	-2.52		
GPI	P06744	0.22	2.63	-0.11		
ZMYM4	Q5VZL5	2.51	2.61	0.02		=
MYL6	P60660	0.32	2.46	-1.73		
C1QBP	Q07021	0.44	2.45	-1.74		
MTA2	O94776	0.48	2.44	-1.40	NuRD complex	↓
RPLP1	P05386	0.35	2.36	-2.18		
MYH9	P35579	0.02	2.35	-2.20		
ACTR3	P61158	-0.20	2.34	-2.67	Arp 2/3 complex	↓
PSMB6	P28072	-0.55	2.30	0.65		

Apéndice IV

Proteína	ID Proteína	Δ27 vs IgG	WT vs IgG	Δ27 vs WT	Componente celular (Gene Ontology)	Cambio MUT vs WT
BUB3	O43684	0.33	2.27	-0.55	Actina	↓
ACTG1	P63261	-0.02	2.27	-2.52		
MYH10	P35580	-0.50	2.24	-2.41		
GATAD2B	Q8WXI9	1.03	2.22	-1.69	NuRD complex	↓
HCLS1	P14317	0.17	2.22	-2.15	Arp 2/3 complex	↓
ARPC1B	O15143	-0.27	2.20	-2.39		
TMOD3	Q9NYL9	-0.15	2.18	-1.80	Arp 2/3 complex	↓
TCEB1	Q15369	0.06	2.18	-2.00		
ATAD3A	Q9NVI7	1.87	2.16	-0.18	Arp 2/3 complex	↓
HSPA8	P11142	1.67	2.15	-0.32		
BTF3	P20290	2.25	2.12	0.24	Arp 2/3 complex	↓
KPNA4	O00629	0.49	2.11	-1.56		
MYCBP	Q99417	1.84	2.10	0.04	Arp 2/3 complex	↓
TCEB2	Q15370	1.91	2.09	-0.07		
MYO1D	O94832	-0.52	2.07	-2.09	Arp 2/3 complex	↓
PRKCB	P05771	0.06	2.07	-1.89		
DBN1	Q16643	-0.23	2.06	-2.83	Arp 2/3 complex	↓
NAA10	P41227	1.77	2.04	-0.16		
CCS	O14618	2.38	2.03	0.46	Arp 2/3 complex	↓
MYL12A	P19105	-0.31	2.02	-1.78		
PDHB	P11177	0.13	1.99	-0.17	Arp 2/3 complex	↓
CORO1C	Q9ULV4	-0.33	1.91	-2.73		
CBX1	P83916	0.06	1.90	-1.72	Arp 2/3 complex	↓
TRMT10C	Q7L0Y3	1.62	1.89	-0.15		
PRPF19	Q9UMS4	0.34	1.89	-1.50	Arp 2/3 complex	↓
NAGK	Q9UJ70	0.60	1.88	-1.30		
RNH1	P13489	0.63	1.88	-0.01	Arp 2/3 complex	↓
ACTR2	P61160	-0.08	1.87	-2.04		
SUPT16H	Q9Y5B9	0.34	1.87	-0.85	Arp 2/3 complex	↓
RPLP2	P05387	0.47	1.83	-1.08		
LYN	P07948	0.00	1.83	-2.72	Arp 2/3 complex	↓
MYL9	P24844	0.00	1.82	-1.83		
ACTB	P60709	0.04	1.81	-1.59	SWI/SNF - Actina	↓
ARPC3	O15145	-0.35	1.80	-2.17	Arp 2/3 complex	↓
CAPZB	P47756	-0.07	1.78	-1.95	Arp 2/3 complex	↓
PDHA1	P08559	0.06	1.78	-1.60		
RRBP1	Q9P2E9	0.06	1.77	-1.59	Arp 2/3 complex	↓
SPTAN1	Q13813	-0.18	1.77	-2.00		
PPP2R2A	P63151	0.39	1.76	-0.46	Arp 2/3 complex	↓
CSNK2B	P67870	-0.06	1.74	-2.12		
RPL12	P30050	0.68	1.74	-1.09	Arp 2/3 complex	↓
RBBP7	Q16576	0.81	1.74	-1.18		
BCAS2	O75934	1.38	1.72	-0.23	Arp 2/3 complex	↓
PML	P29590	2.07	1.72	0.47		
BST2	Q10589	0.33	1.72	-2.38	Arp 2/3 complex	↓
PRPSAP2	O60256	1.62	1.71	0.02		
NPM1	P06748	0.82	1.70	-0.93	Arp 2/3 complex	↓
AP3B1	O00203	0.44	1.70	-1.05		
LSP1	P33241	0.32	1.69	-1.64	Arp 2/3 complex	↓
TLN1	Q9Y490	1.36	1.68	-0.21		
MCM7	P33993	1.12	1.67	-0.28	Arp 2/3 complex	↓
RPL5	P46777	0.73	1.66	-0.80		

Proteína	ID Proteína	Δ27 vs IgG	WT vs IgG	Δ27 vs WT	Componente celular (Gene Ontology)	Cambio MUT vs WT
TPM4	P67936	0.40	1.65	-1.07	Arp 2/3 complex	↓
RPL3	P39023	0.57	1.63	-1.00		
SUGT1	Q9Y2Z0	0.67	1.63	-0.61		
KPNA1	P52294	0.90	1.61	-0.59		
RPL27A	P46776	0.22	1.58	-0.18		
MYL1	P05976	-0.54	1.57	-1.50		
ARPC5	O15511	-0.33	1.57	-1.86		
CAPZA1	P52907	-0.31	1.57	-2.04		
SEC23B	Q15437	-0.80	1.55	-1.37		
GRB2	P62993	0.10	1.54	-1.45		
ARPC2	O15144	0.23	1.53	-1.80	Arp 2/3 complex	↓
IKZF3	Q9UKT9	0.06	1.53	-2.57	SWI/SNF	↓
RBM14	Q96PK6	0.71	1.51	-0.67		
MTPN	P58546	0.52	1.51	-0.49		
COX5B	P10606	1.38	1.50	0.16		
ATP6V1A	P38606	-0.26	1.50	-0.62		
DPF1	Q92782	0.06	1.40	-1.22		
HSPA5	P11021	1.60	1.47	0.41		
GEMIN6	Q8WXD5	1.81	1.47	0.57		
PCBP1	Q15365	3.21	0.58	0.09		
LRRC15	Q8TF66	2.74	0.35	4.30		
HVCN1	Q96D96	2.61	0.70	-0.48		
PRPS1	P60891	2.20	-0.22	0.46		
SON	P18583	2.25	0.55	0.28		
ERLIN2	O94905	2.10	0.56	1.58		
DSG4	Q86SJ6	1.92	0.35	3.69		
HNRNPA0	Q13151	1.78	0.35	1.55		
PNN	Q9H307	1.76	1.22	-0.64		
CYB5R3	P00387	1.66	0.14	1.20		
DDX39A	O00148	1.65	0.59	0.78		
APOD	P05090	1.55	-0.40	0.65		

APÉNDICE V. Artículo de Tesis y artículos adicionales publicados durante la Tesis doctoral

Leukemia (2020) 34:2722–2735
<https://doi.org/10.1038/s41375-020-0919-5>

ARTICLE



Lymphoma

Frequent mutations in the amino-terminal domain of BCL7A impair its tumor suppressor role in DLBCL

Carlos Balañas-Gavira^{1,2,3} · María I. Rodríguez^{1,2,4} · Alvaro Andrades^{1,2} · Marta Cuadros^{1,4,5} · Juan Carlos Álvarez-Pérez^{1,2} · Ángel F. Álvarez-Prado⁶ · Virginia G. de Yébenes^{6,7} · Sabina Sánchez-Hernández⁸ · Elvira Fernández-Vigo⁹ · Javier Muñoz⁹ · Francisco Martín⁸ · Almudena R. Ramiro⁶ · José A. Martínez-Climent¹⁰ · Pedro P. Medina^{1,2,4}

Received: 2 October 2019 / Revised: 6 June 2020 / Accepted: 9 June 2020 / Published online: 24 June 2020
 © The Author(s), under exclusive licence to Springer Nature Limited 2020

Abstract

Mutations in genes encoding subunits of the SWI/SNF chromatin remodeling complex are frequently found in different human cancers. While the tumor suppressor function of this complex is widely established in solid tumors, its role in hematologic malignancies is largely unknown. Recurrent point mutations in *BCL7A* gene, encoding a subunit of the SWI/SNF complex, have been reported in diffuse large B-cell lymphoma (DLBCL), but their functional impact remains to be elucidated. Here we show that *BCL7A* often undergoes biallelic inactivation, including a previously unnoticed mutational hotspot in the splice donor site of intron one. The splice site mutations render a truncated *BCL7A* protein, lacking a portion of the amino-terminal domain. Moreover, restoration of wild-type *BCL7A* expression elicits a tumor suppressor-like phenotype in vitro and in vivo. In contrast, splice site mutations block the tumor suppressor function of *BCL7A* by preventing its binding to the SWI/SNF complex. We also show that *BCL7A* restoration induces transcriptomic changes in genes involved in B-cell activation. In addition, we report that SWI/SNF complex subunits harbor mutations in more than half of patients with germinal center B-cell (GCB)-DLBCL. Overall, this work demonstrates the tumor suppressor function of *BCL7A* in DLBCL, and highlights that the SWI/SNF complex plays a relevant role in DLBCL pathogenesis.

Supplementary information The online version of this article (<https://doi.org/10.1038/s41375-020-0919-5>) contains supplementary material, which is available to authorized users.

✉ Pedro P. Medina
 pedromedina@ugr.es

¹ Gene Expression Regulation and Cancer Group (CTS-993). GENYO. Centre for Genomics and Oncological Research: Pfizer-University of Granada-Andalusian Regional Government, Granada, Spain

² Department of Biochemistry and Molecular Biology I, University of Granada, Granada, Spain

³ FIBAO - Fundación Pública para la Investigación Biosanitaria de Andalucía Oriental - Alejandro Otero, Granada, Spain

⁴ Instituto de Investigación Biosanitaria de Granada (ibis). GRANADA), Granada, Spain

⁵ Department of Biochemistry and Molecular Biology III and Immunology, University of Granada, Granada, Spain

Introduction

Diffuse large B-cell lymphoma (DLBCL) is the most common lymphoma, with an annual incidence of over 100,000 cases worldwide. Although most DLBCL patients achieve complete remission with frontline R-CHOP immunochemotherapy,

⁶ B Cell Biology Lab, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain

⁷ Department of Immunology, Ophthalmology and ENT, School of Medicine, Complutense University, 12 de Octubre Health Research Institute (imas12), Madrid, Spain

⁸ Genomic Medicine Department, GENYO, Centre for Genomics and Oncological Research, Pfizer-University of Granada-Andalusian Regional Government, Granada, Spain

⁹ Proteomics Unit, CNIO—Spanish National Cancer Research Center, Madrid, Spain

¹⁰ Division of Haemato-Oncology, Centre for Applied Medical Research (CIMA), University of Navarra, CIBERONC, IDISNA, Pamplona, Spain

SPRINGER NATURE

Artículos adicionales publicados durante la Tesis doctoral:

Arenas, A.M., Cuadros, M., Andrades, A., Garcia, D.J., Coira, I.F., Rodriguez, M.I., **Baliñas-Gavira, C.**, Peinado, P., Alvarez-Perez, J.C., and Medina, P.P. (2020). LncRNA DLG2-AS1 as a Novel Biomarker in Lung Adenocarcinoma. Cancers 12(8).

Martin-Padron, J., Boyero, L., Rodriguez, M.I., Andrades, A., Diaz-Cano, I., Peinado, P., **Baliñas-Gavira, C.**, Alvarez-Perez, J.C., Coira, I.F., Farez-Vidal, M.E and Medina, P.P. (2020). Plakophilin 1 enhances MYC translation, promoting squamous cell lung cancer. Oncogene 39, 5479-5493.

Cuadros, M., Andrades, A., Coira, I.F., **Baliñas, C.**, Rodriguez, M.I., Alvarez-Perez, J.C., Peinado, P., Arenas, A.M., Garcia, D.J., Jimenez, P and Medina, P.P. (2019). Expression of the long non-coding RNA TCL6 is associated with clinical outcome in pediatric B-cell acute lymphoblastic leukemia. Blood Cancer J 9, 93.

Cuadros, M., Sanchez-Martin, V., Herrera, A., **Baliñas, C.**, Martin-Padron, J., Boyero, L., Peinado, P., and Medina, P.P. (2017). BRG1 regulation by miR-155 in human leukemia and lymphoma cell lines. Clin Transl Oncol 19, 1010-1017.

Schiaffino-Ortega, S., **Balinas, C.**, Cuadros, M., and Medina, P.P. (2014). SWI/SNF proteins as targets in cancer therapy. J Hematol Oncol 7, 81.

