





# Neoplasias linfoides, célula de origen e identidad tumoral

**Idoia García Ramírez**

Instituto de Biología Molecular y Celular del Cáncer

Universidad de Salamanca-CSIC

**Salamanca, 2017**

El trabajo recogido en la presente tesis doctoral ha dado lugar a la elaboración de 2 artículos científicos:

*Crebbp loss cooperates with Bcl2 overexpression to promote lymphoma in mice.* **García-Ramírez I**, Tadros S, González-Herrero I, Martín-Lorenzo A, Rodríguez-Hernández G, Moore D, Ruiz-Roca L, Blanco O, Alonso-López D, Rivas JL, Hartert K, Duval R, Klinkebiel D, Bast M, Vose J, Lunning M, Fu K, Greiner T, Rodrigues-Lima F, Jiménez R, Criado FJG, Cenador MBG, Brindle P, Vicente-Dueñas C, Alizadeh A, Sánchez-García I\*, Green MR\*. **Blood**. 2017 May 11; 129 (19): 2645-2656. doi: 10.1182/blood-2016-08-733469. Epub 2017 Mar 13.

*Transient expression of Lmo2 is sufficient for oncogenic function and induction of T-cell leukemia.* **Idoia García-Ramírez\***, Sanil Bhatia\*, Alberto Martín-Lorenzo, Guillermo Rodríguez-Hernández, Inés González-Herrero, Salma Parvin, Oskar Haas, Wilhelm Woessmann, Martin Stanulla, Martin Schrappe, Yasodha Natkunam, Alberto Orfao, Verónica Domínguez, Belén Pintado, Oscar Blanco, Diego Alonso-López, Javier De Las Rivas, Rafael Jiménez, Francisco Javier García Criado, María Begoña García Cenador, Carolina Vicente-Dueñas, Izidore S. Lossos, Arndt Borkhardt<sup>#</sup>, Isidro Sánchez-García<sup>#</sup> and Julia Hauer<sup>#</sup>. **THE EMBO JOURNAL**. SUBMITTED.

D. **Isidro Sánchez García**, Investigador Científico del Consejo Superior de Investigaciones Científicas (C.S.I.C) en el Instituto de Biología Molecular y Celular del Cáncer (IBMCC) de Salamanca.

CERTIFICA:

Que el trabajo titulado “**Neoplasias linfoides, célula de origen e identidad tumoral**”, que presenta la doctoranda **Dña. Idoia García Ramírez**, ha sido realizado bajo su dirección en el Instituto de Biología Molecular y Celular del Cáncer, y reúne, a su juicio, originalidad y contenidos suficientes para ser presentado ante el correspondiente Tribunal y optar al Título de Doctor por la Universidad de Salamanca.

Y para que así conste, y a los efectos oportunos, expiden el presente certificado en Salamanca a 2 de Noviembre de 2017.



The image shows a handwritten signature in blue ink, which is a stylized, cursive representation of the name 'Isidro Sánchez García'. To the right of the signature is a circular blue stamp. The stamp contains the text 'CENTRO DE INVESTIGACIÓN DEL CÁNCER' around the top inner edge, 'USAL - CSIC' in the center, and 'IBMCC' at the bottom inner edge.

Fdo. **I. Sánchez García**

La investigación en nuestro laboratorio está financiada parcialmente por FEDER y por MINECO (SAF2012-32810, SAF2015-64420-R, Y Red de Excelencia Consolidar OncoBIO SAF2014-57791-REDC), por el Instituto de Salud Carlos III (PIE14/00066), por el ISCIII-Plan de ayudas IBSAL 2015 Proyectos Integrados (IBY15/00003), por la Junta de Castilla y León (BIO/SA51/15, CSI001I14, UIC-017, y CSI001U16), por la Fundación Inocente Inocente, por el proyecto ARIMMORA (European Union's Seventh Framework Programme (FP7/2007-2013)) y por la Fundación Carreras Alemana (DJCLS R13/26). Nuestro laboratorio es miembro del EuroSyStem y la Red DECIDE, financiadas por el programa FP7 de la Unión Europea.

Durante la realización de este proyecto doctoral he disfrutado de un contrato predoctoral para la formación de doctores del Subprograma Estatal de Formación del Ministerio de Economía y Competitividad con referencia BES-2013-063789 y de una beca con cargo a la investigación (Contrato Art.83 LOU (GRIFOLS-USAL)) con referencia LBNL.

En ocasiones la vida te enseña a darte cuenta de las cosas realmente importantes, y a mí este último año me ha enseñado que una de estas cosas es dar las gracias y no rendirse nunca. Por eso, me gustaría aprovechar estas líneas para dar las gracias a todas las personas que han hecho posible que este trabajo sea una realidad.

En primer lugar, quisiera agradecer a mi director de tesis su apoyo durante este periodo, dado que sin él todo este trabajo no habría sido posible. Gracias Isidro por enseñarme en estos años el valor del esfuerzo y el sacrificio, y la importancia de no rendirse nunca.

A todos los miembros del laboratorio 13: Alberto, Guille, Inés, Caro, Ana, Xiomara y Sara. Y a todos aquellos que formaron parte de él: Marcos, Lucía, Isa y especialmente a la memoria de Ester. Gracias por vuestra ayuda y comprensión permanentes. Gracias por todo lo que me habéis enseñado en estos años.

Al Dr. Jesús Pérez Losada y a todos los miembros de su laboratorio, Adri, Aurora, Milena y Roberto gracias por vuestra ayuda y por compartir semanalmente vuestros conocimientos en nuestros seminarios.

Al Profesor Rafael Jiménez del Dto. De Fisiología y Farmacología de la Universidad de Salamanca por su constante apoyo y sus buenos consejos.

Al Dr. Dionisio Martín Zanca del Instituto de Biología Funcional y Genómica de Salamanca, por su asesoramiento científico en los seminarios semanales.

Al Dr. César Coboleda Hernández y a Elena Campos Sánchez del Centro de Biología Molecular de Madrid, por toda la ayuda brindada durante este tiempo y por los buenos consejos aportados.

Al Profesor Francisco Javier García Criado, Decano de la Facultad de Medicina de la Universidad de Salamanca, por su ayuda incondicional al grupo.

A la Dra. Begoña García Cenador, del Dto. De Cirugía de la Universidad de Salamanca, por estar disponible para ayudarnos siempre que lo necesitamos y por su cercanía y simpatía.

A todo el personal de apoyo del Instituto de Biología Molecular y Celular del Cáncer de Salamanca: mantenimiento, informática, microscopia, cocina...

A todo el personal de gestión del Instituto de Biología Molecular y Celular del Cáncer, especialmente a Javier Beltrán por tu ayuda en lo profesional y lo personal.

A todo el personal del Servicio de Experimentación Animal de la Universidad de Salamanca, en especial a Paco, Josito, Jorge, Sara y Juan por hacer siempre todo lo posible por facilitarnos el trabajo.

Al Servicio de Citometría de la Universidad de Salamanca dirigido por el Dr. Alberto Orfao, especialmente a M<sup>a</sup> Luz, por ayudarme siempre en el desarrollo de los experimentos de citometría.

Al Servicio de Genómica del Instituto de Biología Molecular y Celular del Cáncer, especialmente a la Dra. Encarna Fermiñán Benito y a la Dra. Eva M<sup>a</sup> García García, por el procesado de las muestras para los microarrays de expresión.

A la Dra. Belén Pintado, directora de la Unidad de Transgénesis del Centro Nacional de Biotecnología de Madrid por su ayuda para la generación de los modelos murinos del presente trabajo.

Al Dr. Diego Alonso López, del Servicio de Bioinformática del Instituto de Biología Molecular y Celular del Cáncer, por su ayuda en el análisis de los microarrays de expresión.

Al servicio de Anatomía Patológica del Hospital Universitario de Salamanca, y muy especialmente a la Dra. Teresa Flores Corral y al Dr. Óscar Blanco que han llevado a cabo el procesado y diagnóstico anatomopatológico de las muestras de ratón.

Al Dr. Michael Green y a los miembros de su equipo en la Universidad de Nebraska, por todo el análisis molecular y epigenético del proyecto Crebbp, sin ellos el trabajo no habría podido ser terminado.

Al Dr. Arndt Borkhardt y a todos los miembros de su equipo en el Hospital de Dusseldorf, en especial a Franz, Julia, Sanil, Uthe y Daniel. Gracias por toda la ayuda científica y personal. Es difícil conseguir que 2 laboratorios de países diferentes funcionen como uno solo, y con vosotros ha sido posible.

Me gustaría dar las gracias también a mis amigos de toda la vida, Mary, Bez y Trini, porque aunque nos separe un océano siempre tendremos un día al año para arreglar el mundo. A Estrella, Chana, Javi e Isa, los 5 empezamos esta andadura y aunque la vida nos ha llevado por diferentes caminos, siempre formareis parte de mi vida.

A los amigos que no tenía y que gracias a esta experiencia he conocido, Javi, Jaime, Pili y muy especialmente a Athe, gracias por estar siempre ahí, por ser mi familia de no sangre, por ser como una hermana.

A Alberto y Guille, como dijo un gran sabio, "yo no vengo aquí a hacer amigos". Gracias por soportarme cada día (se que no ha sido fácil), por vuestra ayuda y por todos los buenos ratos que he pasado y espero seguir pasando a vuestro lado.

Pero sobre todo, quiero agradecer su apoyo y amor incondicional a mis abuelos, tíos, suegros y, especialmente, a mis padres, Miguel y Juani. Se que no ha sido un año fácil pero siempre habéis estado ahí, apoyándome y haciendo que sea la persona que soy hoy.

Y por último dar las gracias a Antonio, por soportar estos 6 años separados como nadie lo habría hecho, por tu paciencia infinita, por tu apoyo, por tu cariño y por siempre saber como sacarme una sonrisa cuando nadie puede hacerlo. Este trabajo es tan tuyo como mío.

En Salamanca, a 2 de Noviembre 2017

Idoia García Ramírez

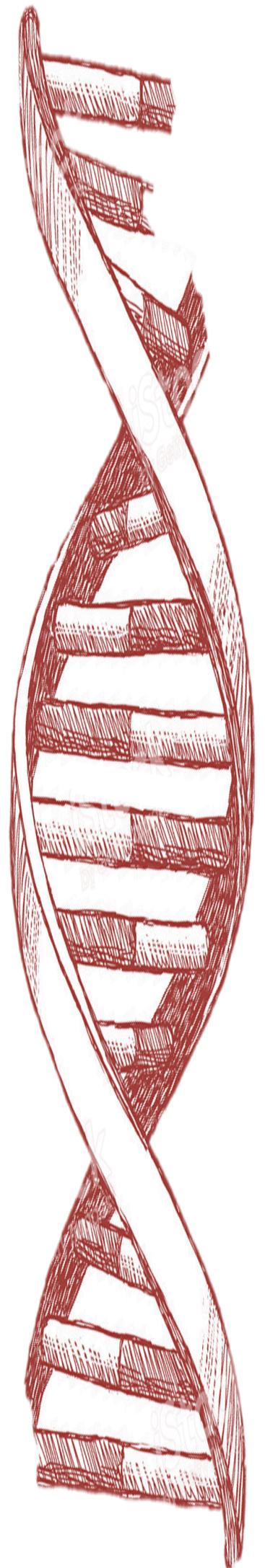
**A mi familia y a Antonio**

*“La ciencia humana consiste más en destruir errores que en descubrir verdades”*

Sócrates, Filósofo griego.



*indice*





## Índice

<b>ABREVIATURAS .....</b>	<b>3</b>
<b>INTRODUCCIÓN.....</b>	<b>13</b>
<b>1. Función de los mecanismos moleculares en el desarrollo del cáncer.....</b>	<b>13</b>
1.1 Células cancerosas y adicción oncogénica.....	14
1.2 Interacción oncogén-célula diana.....	16
<b>2. El concepto de “Célula madre del cáncer o <i>Cancer Stem Cell</i>” .....</b>	<b>18</b>
2.1 La Teoría de la Célula Madre del Cáncer: antecedentes, evidencias e hipótesis.....	18
2.2 Implicaciones prácticas de la Teoría de la Célula Madre del Cáncer.....	20
2.3 Célula de origen del cáncer vs célula madre del cáncer.....	23
<b>3. El mecanismo de reprogramación tumoral como origen del cáncer. ....</b>	<b>25</b>
3.1 La reprogramación como origen del cáncer.....	25
3.2 Modelos experimentales in vivo de reprogramación tumoral de células madre.....	28
<b>4. <i>CREBBP</i> y su función en el desarrollo del linfoma folicular y del linfoma difuso de células B grandes.....</b>	<b>29</b>
4.1 Los linfomas no Hodgkin: definición, epidemiología, clínica y clasificación.....	29
4.1.1 El linfoma folicular: definición, epidemiología, inmunofenotipo, clínica, y etiología.....	31
4.1.2 El linfoma difuso de células B grandes: definición, epidemiología, clasificación, inmunofenotipo, clínica, y etiología.....	33
4.2 La proteína <i>CREBBP</i> humana: descubrimiento, estructura, función y.....	36
patogénesis.....	36
<b>5. <i>LMO2</i> y su función en el desarrollo de la leucemia linfoblástica aguda de células T.....</b>	<b>40</b>
5.1 Leucemia linfoblástica aguda de células T: definición, epidemiología, clínica, clasificación.....	40
5.2 La proteína <i>LMO2</i> humana: descubrimiento, estructura, función y.....	44
patogénesis.....	44
5.3 <i>LMO2</i> y terapia génica.....	50

<b>HIPÓTESIS DE TRABAJO.....</b>	<b>55</b>
<b>OBJETIVOS.....</b>	<b>60</b>
<b>RESULTADOS .....</b>	<b>65</b>
<b>1. La pérdida de <i>Crebbp</i> a nivel de células pro-B coopera con la sobreexpresión de <i>Bcl2</i> para promover la linfomagénesis.....</b>	<b>65</b>
1.1 La pérdida de <i>Crebbp</i> a nivel de célula pro-B produce una alteración del desarrollo normal de las células B.....	65
1.2 La pérdida de <i>Crebbp</i> combinada con la sobreexpresión de <i>Bcl2</i> en células pro-B conduce al desarrollo de linfomas de células B. ....	71
1.3 La inactivación de <i>Crebbp</i> en células B induce de forma indirecta la expresión de <i>Myc</i> .....	77
1.4 La inactivación de <i>Crebbp</i> está relacionada con una firma epigenética asociada a <i>Myc</i> . ....	80
1.5 Presencia de diferentes patrones mutacionales de <i>CREBBP</i> en el linfoma folicular en comparación con el linfoma difuso de células B grandes.....	83
<b>2. La expresión transitoria de <i>Lmo2</i> en células madre hematopoyéticas (HSCs) causa T-ALL dependiente del timo.....</b>	<b>86</b>
2.1 Generación de una cepa de ratón en la que la expresión condicional de <i>Lmo2</i> está dirigida a HSCs. ....	86
2.2 La expresión de <i>Lmo2</i> en células madre hematopoyéticas de ratón causa exclusivamente T-ALL, similar a la T-ALL humana.....	88
2.3 <i>Lmo2</i> funciona como un oncogén “hit-and-run” en el desarrollo de la T-ALL. ....	95
2.4 Las células T leucémicas de los ratones <i>Sca1-Lmo2</i> muestran alteraciones genéticas y moleculares análogas a las que aparecen en las T-ALL humanas. ....	100
2.5 Las T-ALL <i>Lmo2</i> <sup>neg</sup> de los ratones <i>Sca1-Lmo2</i> comparten eventos secundarios críticos con las T-ALL humanas y las T-ALL de los ratones <i>Rosa26-Lmo2+Sca1-Cre</i> .....	102
2.6 La adquisición de las alteraciones genéticas secundarias tiene lugar en el timo durante el desarrollo de las T-ALL. ....	104
<b>DISCUSIÓN .....</b>	<b>113</b>

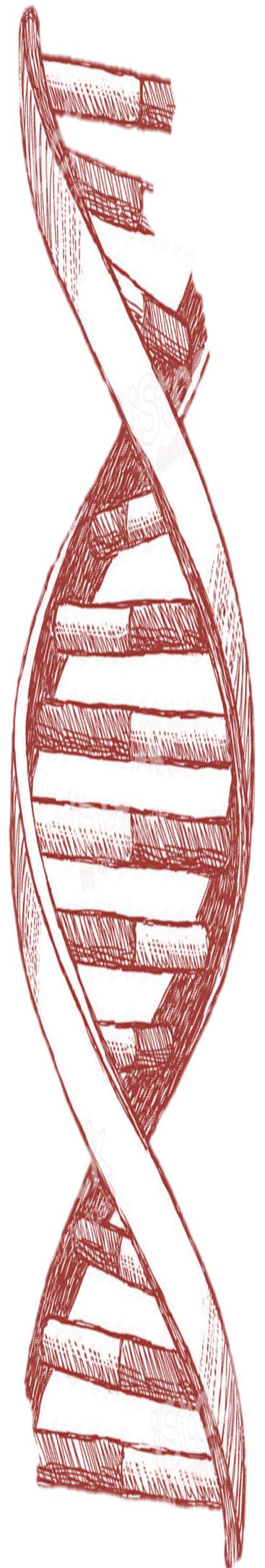
1. La pérdida de <i>Crebbp</i> a nivel de células pro-B coopera con la sobreexpresión de <i>Bcl2</i> para promover la linfomagénesis.....	113
2. Presencia de diferentes patrones mutacionales de <i>CREBBP</i> en el linfoma folicular en comparación con el linfoma difuso de células B grandes. ....	116
3. La expresión transitoria de <i>Lmo2</i> en células madre hematopoyéticas (HSCs) causa T-ALL dependiente del timo.....	118
<b>CONCLUSIONES.....</b>	<b>126</b>
<b>MATERIAL Y MÉTODOS .....</b>	<b>131</b>
1. Animales utilizados en el estudio.....	131
2. Genotipado de los ratones mediante <i>Southern blot</i> . ....	132
3. Genotipado de los ratones mediante PCR. ....	133
3.1 PCR de GFP. ....	133
3.2 PCR de neo. ....	133
3.3 PCR de <i>Crebbp</i> silvestre (WT) y condicional.....	134
3.4 PCR de mb1-cre. ....	134
3.5 PCR Emu- <i>Bcl2</i> . ....	135
4. Análisis de citometría de flujo.....	136
5. Estudio de formación del centro germinal.....	136
6. Análisis histológicos e inmunohistoquímica (IHC). ....	137
7. Análisis de expresión de proteína mediante <i>Western blot</i> . ....	137
8. Extracción de ADN .....	138
9. Análisis de clonalidad de las inmunoglobulinas (VDJ). ....	138
10. PCR del transgén <i>Crebbp</i> delecionado.....	139
11. Análisis de clonalidad del gen del receptor de linfocitos T (TCR).....	139
12. Hibridación genómica comparada o CGH. ....	140
13. Extracción de ARN.....	140
14. Perfil de expresión génica .....	141
15. Análisis de enriquecimiento.....	141
16. Secuenciación mediante inmunoprecipitación de la cromatina ( <i>ChIP-seq</i> ) .....	142
17. Análisis de enriquecimiento de los motivos de ADN.....	143
18. Análisis de hibridación fluorescente <i>in situ</i> (FISH). ....	143

<b>19. Preparación de librerías del exoma completo humano y de ratón (WES).</b>	<b>144</b>
.....	144
19.1 Secuenciación de nueva generación (NGS) de los linfomas de células B.	144
.....	144
19.2 Secuenciación de nueva generación (NGS) de las T-ALL humanas y murinas.....	145
<b>ANEXOS.....</b>	<b>151</b>
<b>BIBLIOGRAFÍA.....</b>	<b>157</b>





*Abreviaturas*





**ABREVIATURAS****A**

<b>ABC-DLBCL:</b>	Linfoma difuso de células B grandes activadas
<b>ALL:</b>	Leucemia linfoblástica aguda
<b>AML:</b>	Leucemia mieloide aguda
<b>AMPc:</b>	Adenosin monofosfato cíclico
<b>ADN:</b>	Ácido desoxirribonucleico
<b>ADNc:</b>	Ácido desoxirribonucleico complementario
<b>ANOVA:</b>	Análisis de la varianza
<b>APC:</b>	Aloficocianina
<b>ARN:</b>	Ácido ribonucleico
<b>ARNi:</b>	Ácido ribonucleico de interferencia
<b>ARNm:</b>	Ácido ribonucleico mensajero

**B**

<b>B-ALL:</b>	Leucemia linfoblástica aguda de células B
---------------	---

**C**

<b>CCO:</b>	Célula de origen del cáncer
<b>CGH:</b>	Hibridación genómica comparada
<b>ChIP-seq:</b>	Secuenciación de inmunoprecipitación de cromatina
<b>CLL:</b>	Leucemia linfocítica crónica

## ABREVIATURAS

---

---

**CML:** Leucemia mieloide crónica

**Cr:** Cromosoma

**CSC:** Célula madre del cáncer

### D

**ddH<sub>2</sub>O:** Agua bidestilada

**DLBCL:** Linfoma difuso de células B grandes

**DLBCL-NOS:** Linfoma difuso de células B grandes no especificado

### E

**EDTA:** Ácido etilendiaminotetraacético

**ESC:** Célula madre embrionaria

**ETP:** Progenitor T temprano

### F

**FACS:** Separación celular por activación por fluorescencia

**FCS:** Suero bovino fetal

**FDR:** Tasa de descubrimiento de falsos positivos

**FISH:** Hibridación fluorescente in situ

**FITC:** Isotiocianato de fluoresceína

**FL:** Linfoma folicular

**G**

<b>GC:</b>	Centro germinal
<b>GCB:</b>	Célula B del centro germinal
<b>GCB-DLBCL:</b>	Linfoma difuso de células B grandes del centro germinal
<b>GSEA:</b>	Análisis de enriquecimiento mediante de grupos de genes
<b>GFP:</b>	Proteína de fluorescencia verde

**H**

<b>H3K18:</b>	Lisina 18 de la histona 3
<b>H3K18Ac:</b>	Acetilación de la lisina 18 de la histona 3
<b>HAT:</b>	Histona acetiltransferasa
<b>HDAC:</b>	Histona deacetilasa
<b>HRP:</b>	Peroxidasa de rábano
<b>HSC:</b>	Célula madre hematopoyética
<b>HSC/PC:</b>	Célula madre/progenitora hematopoyética

**I**

<b>Ig:</b>	Inmunoglobulina
<b>IL2RG:</b>	Receptor <i>gamma</i> de la Interleucina 2
<b>Indel:</b>	Inserción/Delección
<b>iPSC:</b>	Célula madre pluripotente inducida

### **K:**

**KAT:** Lisina acetiltransferasa

**kb:** Kilobases

### **M**

**MACS:** Separación celular por activación magnética

**MALT:** Tejido linfóide asociado a mucosas

**MAPK:** Proteínas quinasas activadas por mitógenos

**Mb:** Megabases

**MHC-II:** Complejo mayor de histocompatibilidad de clase II

**MPW:** Agua Millipore

### **N**

**NGS:** Secuenciación de nueva generación

**NHL:** Linfoma no Hodgkin

**NOD/SCID:** Diabéticos no obesos con inmunodeficiencia severa combinada

### **O**

**OMS:** Organización Mundial de la Salud

### **P**

**pb:** Pares de bases

<b>PBS:</b>	Tampón fosfato salino
<b>PCR:</b>	Reacción en cadena de la polimerasa
<b>PE:</b>	Ficoeritrina
<b>PerCP:</b>	Peridina clorofila
<b>PI:</b>	Ioduro de propidio
<b>PNA:</b>	Aglutinina de cacahuete
<b>PTNFL:</b>	Linfoma folicular ganglionar de tipo pediátrico

## **Q**

<b>qPCR:</b>	Reacción en cadena de la polimerasa cuantitativa
--------------	--

## **R**

<b>RBCL:</b>	Tampón de lisis de eritrocitos
<b>RMA:</b>	<i>Robust Multichip Average</i>

## **S**

<b>SAM:</b>	<i>Significance Analysis of Microarrays</i>
<b>SCID-X1:</b>	Síndrome de inmunodeficiencia combinada severa ligada al cromosoma X
<b>SDS:</b>	Dodecil sulfato sódico
<b>SDS-PAGE:</b>	Electroforesis en gel de poliacrilamida con dodecil sulfato sódico
<b>SHM:</b>	Hipermutación somática
<b>SNC:</b>	Sistema nervioso central

## ABREVIATURAS

---

---

<b>SNP:</b>	Polimorfismo de un solo nucleótido
<b>SNV:</b>	Variación de un solo nucleótido
<b>SRBC:</b>	Solución de eritrocitos de cordero
<b>SRT:</b>	Síndrome de Rubinstein-Taybi
<b>SSC:</b>	Citrato salino de sodio

### T

<b>T-ALL:</b>	Leucemia linfoblástica aguda de células T
<b>TCR:</b>	Receptor de linfocitos T
<b>TdT:</b>	Desoxinucleotidil transferasa terminal
<b>TSS:</b>	Sitio de inicio de la transcripción

### U

<b>UV:</b>	Ultravioleta
------------	--------------

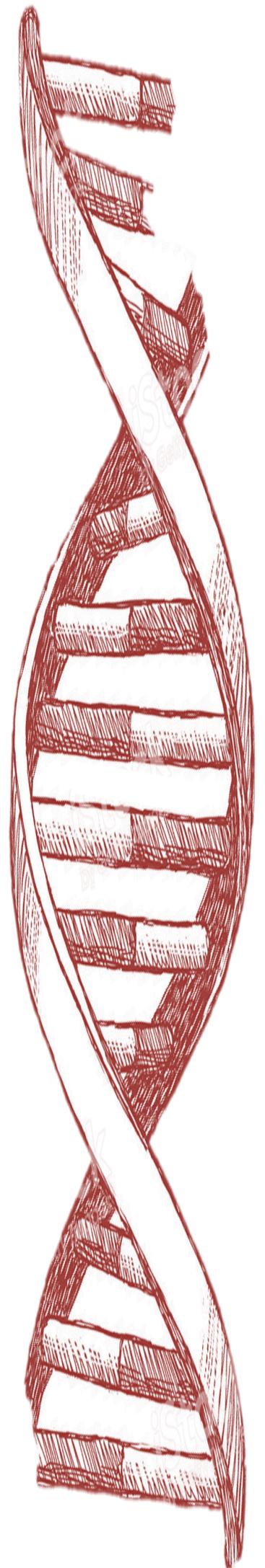
### W

<b>WES:</b>	Secuenciación del exoma completo
<b>WT:</b>	Tipo silvestre





# *Introducción*



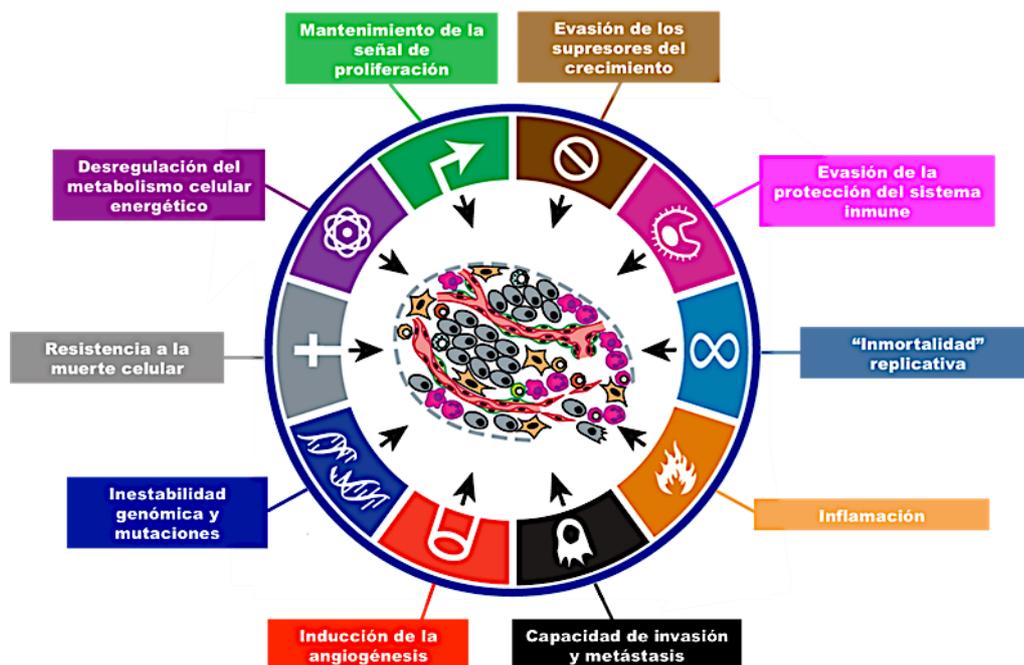


## INTRODUCCIÓN

### 1. Función de los mecanismos moleculares en el desarrollo del cáncer.

Durante décadas, la investigación contemporánea sobre el cáncer ha centrado todos sus esfuerzos en controlar la proliferación anormal de las células tumorales. Este hecho se ha visto reflejado en los enfoques terapéuticos empleados en la clínica para tratar a los pacientes: con muy pocas excepciones, los tratamientos actuales contra el cáncer suelen tener como diana los mecanismos de crecimiento tumoral anormal. Sin embargo, con las terapias actuales, el cáncer sigue siendo una de las principales causas de morbilidad y mortalidad. Concretamente, en 2012 se registraron alrededor de 14 millones de nuevos casos de cáncer en el mundo<sup>1</sup>, siendo la segunda causa de muerte. En 2015 se atribuyeron a esta enfermedad 8,8 millones de defunciones (según datos facilitados por la OMS). Estos datos tan pesimistas son consecuencia, fundamentalmente, del fracaso de las terapias, que a menudo va acompañado por el desarrollo de resistencias a los fármacos y por la diseminación metastásica. Por esta razón, un objetivo prioritario de la investigación frente al cáncer debe ser entender, en primer lugar, cómo contrarrestar los mecanismos que subyacen a la capacidad de las células normales para convertirse en células cancerosas.

La complejidad de las células cancerosas fue resumida, y posteriormente revisada, por Hanahan y Weinberg<sup>2,3</sup> en “nueve alteraciones esenciales en la fisiología celular que rigen colectivamente el crecimiento maligno” (**Figura 1**).



**Figura 1: Principales características del cáncer.**

Adaptada de D. Hanahan *et al.*; Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, (2011)<sup>3</sup>.

Las células cancerosas son la base del cáncer, lo inician y permiten la progresión de la enfermedad. Además, son estas células las portadoras de mutaciones oncogénicas y de genes supresores de tumores, lo que hace que el cáncer se defina como una enfermedad genética<sup>3</sup>. Sin embargo, aún no entendemos suficientemente bien los mecanismos subyacentes que conducen al origen de estas células cancerosas, lo que ha hecho que no se haya conseguido aumentar de manera significativa el porcentaje de pacientes curados ni cambiar la base de los tratamientos usados frente al cáncer<sup>4</sup>. El único factor que se ha reducido considerablemente en los últimos años ha sido la mortalidad global por cáncer, y no por el éxito de los tratamientos, sino por la mejora en las técnicas de detección precoz de la enfermedad<sup>5,6</sup>. Sin embargo, tras un período de remisión, un elevado porcentaje de pacientes recaen.

Por lo tanto, una comprensión completa del proceso del cáncer requiere un conocimiento más detallado de los mecanismos que dan lugar al crecimiento neoplásico, y éste es un requisito previo no sólo para la comprensión de la génesis del cáncer humano, sino también para la identificación de los eventos moleculares responsables del mantenimiento de la enfermedad. A pesar de ello, todos los aspectos relacionados con los mecanismos reguladores del desarrollo en la carcinogénesis han recibido relativamente poca atención a lo largo del proceso de identificación y definición de las características esenciales de las células cancerosas. De hecho, si como se sostiene el destino celular fuese inamovible, entonces el cáncer no se produciría, ya que no se podrían generar otros linajes distintos a los normales, es decir, a los fisiológicos. Es aquí donde los mecanismos que regulan la identidad celular del tumor juegan un importante papel en permitir que el cáncer aparezca, sugiriendo que dichos mecanismos pudieran ser la llave para la eliminación del mismo<sup>7</sup>.

Todo ello hace evidente la necesidad de identificar los mecanismos moleculares responsables del establecimiento de la identidad de las células tumorales y cómo una mejor comprensión de los mismos, debería conducirnos a un conocimiento más profundo de la carcinogénesis y al desarrollo de tratamientos efectivos.

### ***1.1 Células cancerosas y adición oncogénica.***

En las últimas décadas, la investigación científica ha demostrado la importancia de los oncogenes en el cáncer humano. Desde el descubrimiento de que los tumores humanos contienen oncogenes activados<sup>8-13</sup>, se han llevado a cabo muchos estudios para comprender su papel causal en el desarrollo del cáncer. Todo este trabajo ha demostrado que la expresión del oncogén no sólo es necesaria para el inicio del

cáncer, sino también para el mantenimiento de la enfermedad, lo que ha hecho que los oncogenes se mantengan como principales objetivos terapéuticos contra el cáncer. Además, en los modelos de ratón donde la expresión del oncogén está dirigida por promotores específicos de tejido, los tumores surgen con una elevada frecuencia, pero remiten cuando el estímulo oncogénico cesa<sup>14-16</sup>, este hecho sugiere que los oncogenes son el talón de Aquiles de la célula cancerosa<sup>17</sup>. Este modelo actual del cáncer sostiene que los oncogenes tienen una función homogénea en todas las células cancerosas del tumor, ya que la breve inactivación de los diferentes oncogenes inductores puede causar la remisión del cáncer en estos modelos. Desafortunadamente, las terapias basadas en este modelo de cáncer no logran erradicar los tumores en humanos. Estas observaciones clínicas sugieren que, en pacientes humanos, la tumorigénesis inducida por el oncogén puede no ser reversible a través de la inactivación única del (de los) defecto(s) genético(s) que inicia el desarrollo del cáncer. Esto nos hace plantear cuáles son los mecanismos de recaída tumoral que permiten a los tumores evolucionar escapando de la dependencia al oncogén<sup>7</sup>.

El fracaso en el tratamiento del cáncer no puede explicarse simplemente por la existencia de células madre del cáncer (CSCs, por sus siglas en inglés), sobre las que más tarde hablaremos, o por la plasticidad celular de los tumores. De hecho, ambos aspectos sólo implican que una población tumoral genéticamente homogénea puede presentarse como fenotípicamente heterogénea, debido a la presencia de células en diferentes estadios de diferenciación<sup>3</sup>. Sin embargo, las observaciones clínicas, antes mencionadas, derivadas del fracaso terapéutico dirigido a humanos, podrían sugerir que a diferencia de lo que se pensaba, los oncogenes tienen un modo de acción que no es homogéneo en el conjunto de la población de células cancerosas. Esto explicaría por qué los diferentes estadios celulares del cáncer muestran una sensibilidad diferente a las terapias dirigidas frente al oncogén. Además, recientes evidencias genéticas en modelos *in vivo* han demostrado que los oncogenes humanos son capaces de reprogramar células madre o precursoras tempranas hacia células tumorales diferenciadas específicas, pero la presencia del oncogén no es necesaria dentro de las células malignas diferenciadas para mantener la enfermedad<sup>18-23</sup>.

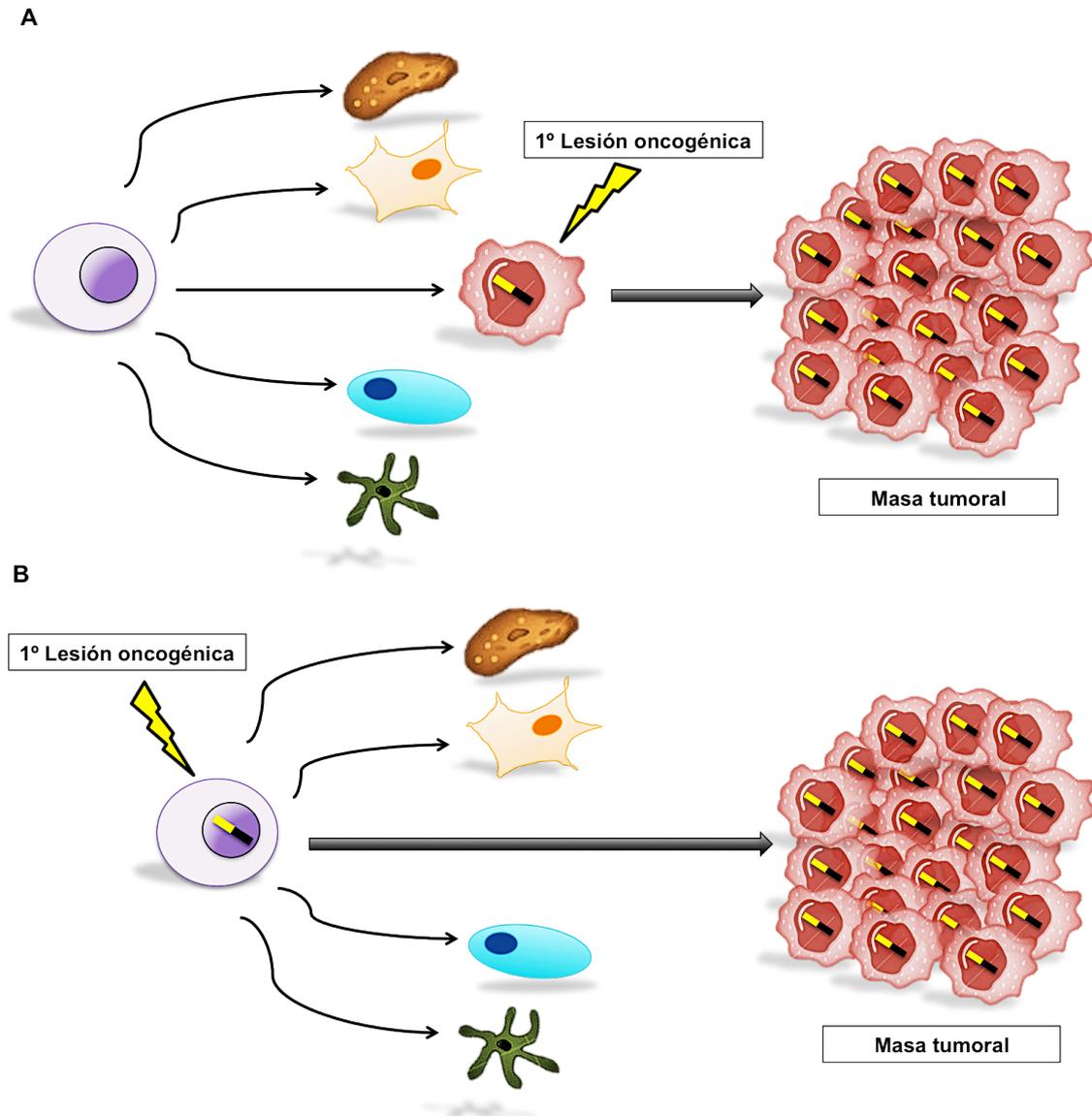
Estos resultados no sólo destacan un papel desconocido hasta el momento para los oncogenes humanos, sino que también ponen de manifiesto que existe un proceso de tumorigénesis previamente no estudiado, en el que la programación del fenotipo maligno ya ha tenido lugar en la etapa de célula madre.

### **1.2 Interacción oncogén-célula diana.**

En los últimos años, ha surgido una nueva corriente que reconoce el papel de la diferenciación aberrante en el origen del cáncer, impulsada principalmente por la aparición de la teoría de las células madre del cáncer. Desde este punto de vista, un conocimiento exhaustivo de los mecanismos de desarrollo mediante los cuales las células diana normales adquieren su identidad tumoral es esencial para comprender el desarrollo del cáncer.

Los estudios moleculares y citogenéticos han identificado que muchos tipos de cáncer están específicamente asociados a eventos genéticos definidos<sup>24</sup>. La expresión de cada una de estas alteraciones genéticas se asocia casi exclusivamente con un subtipo específico de cáncer humano. Estas alteraciones genéticas no sólo tienen importancia como marcadores diagnósticos en la práctica clínica, sino que también proporcionan datos importantes para la comprensión de los mecanismos celulares que se encuentran detrás del desarrollo del cáncer. Sin embargo, estas correlaciones genotipo-fenotipo, establecidas en el cáncer humano, han hecho necesario estudiar la naturaleza de esta estrecha asociación entre cada lesión genética y el fenotipo (el tipo particular de cáncer) con el que se asocia. Como consecuencia de ello, se han planteado dos hipótesis diferentes para explicar dicha correlación. En la hipótesis clásica del inicio y progresión del cáncer, la alteración genética iniciadora tiene lugar en una célula diana comprometida o diferenciada, y esta alteración genética es necesaria para la inmortalización de la célula diana. Dicha célula adquiriría posteriormente lesiones genéticas adicionales a lo largo del tiempo. La adquisición de estos eventos genéticos agravaría la desregulación del comportamiento de la célula diana diferenciada, conduciendo finalmente al desarrollo del cáncer y sus características clínicas (**Figura 2A**). Este modelo es el que se ha asumido tradicionalmente a la hora de estudiar la oncogénesis, dando por supuesto que el fenotipo de las células tumorales es un reflejo de la célula normal diferenciada que dio lugar al tumor en un primer momento. Sin embargo, hace 40 años, se puso en entredicho que esta teoría fuera aplicable a todos los tipos de cáncer. Un ejemplo es el de la leucemia mieloide crónica (CML), sobre la que *Fialkow* y colaboradores sugirieron por primera vez que la enfermedad surgía a partir de una célula madre hematopoyética (HSC) rara transformada, dado que la translocación cromosómica t(9, 22) se podía encontrar en la mayoría de subtipos de células hematopoyéticas diferenciadas<sup>25</sup>. A pesar de ello, en la mayoría de los casos, las células cancerosas comparten similitudes con algunos tipos de células diferenciadas no patogénicas. De

ahí que hoy en día, para cada tipo de cáncer, se asuma que la célula de origen es la correspondiente célula normal diferenciada.



**Figura 2: Modelos propuestos para el papel de los defectos genéticos del cáncer humano en la determinación de la identidad celular del tumor. A)** Modelo clásico, en el que se sostiene que las alteraciones genéticas que aparecen en el cáncer humano actúan sobre células ya comprometidas en un programa de diferenciación específico, de tal manera que el fenotipo tumoral deriva de la célula diana diferenciada inicial. **B)** Modelo alternativo, los últimos descubrimientos apoyan la idea de que la lesión oncogénica actuaría en una célula madre/progenitora imponiendo un destino celular específico, siendo éste específico de cada oncogén y dando lugar a una masa tumoral con una identidad celular diferenciada.

Sin embargo, en los últimos años ha surgido una nueva hipótesis para explicar las correlaciones genotipo-fenotipo observadas entre lesiones genéticas y un tipo tumoral específico. Esta hipótesis plantea la posibilidad de que el oncogén sea el responsable

directo de imponer las características específicas del fenotipo tumoral<sup>7</sup> (**Figura 2B**). Esto es, de hecho, lo que ocurriría en la CML, donde el oncogén se expresa en las células madre hematopoyéticas, pero el fenotipo tumoral se manifiesta como una expansión granulocítica. Además, esta nueva interpretación de la oncogénesis estaría apoyada por la teoría de la célula madre del cáncer (CSC)<sup>4,26</sup>.

## 2. El concepto de “Célula madre del cáncer o *Cancer Stem Cell*”

### **2.1 La Teoría de la Célula Madre del Cáncer: antecedentes, evidencias e hipótesis.**

El hecho de que la mayoría de terapias actuales hayan fracasado a la hora de erradicar el cáncer, ha favorecido la aparición de una nueva hipótesis orientada a entender la biología del cáncer, la de la “Célula Madre del Cáncer” (en inglés “*Cancer Stem Cell*” ó CSC). Esta hipótesis se basa en la existencia de una célula madre del cáncer la cual es capaz de dirigir y mantener el desarrollo del tumor<sup>5,27,28</sup>.

El concepto de CSC se remonta al siglo XIX: trescientos años después de la invención del microscopio, simultáneamente con el surgimiento de la teoría Darwiniana, el fisiólogo alemán Johannes Müller y sus colaboradores aplicaron la microscopía a muestras de tumores humanos. Aplicando los métodos utilizados por los botánicos y los fisiólogos vegetales, Müller transformó la patología y la medicina moderna con su monografía sobre el cáncer. En ella, Müller propuso por primera vez que los tumores estaban compuestos por células “nuevas” dentro del órgano enfermo. Fue un estudiante de Müller, el famoso patólogo Rudolf Virchow, quien años más tarde determinó que todos los tumores derivan de las células normales. Müller y Virchow transformaron la medicina moderna, no sólo inventando el campo de la patología, sino también registrando algunas de las primeras pruebas de que los tumores son heterogéneos<sup>29</sup>.

En 1937, Jacob Furth y colaboradores demostraron que una sola célula leucémica de ratón era capaz de inducir la enfermedad en otro ratón<sup>30</sup>. Sin embargo, el concepto de célula madre del cáncer, como tal, fue propuesto por primera vez hace más de cuatro décadas, para denominar a una fracción de células que aparecían en el mieloma múltiple y que eran capaces de generar nuevas colonias en cultivo<sup>31, 32</sup>.

Más tarde, en 1997, fue demostrada por primera vez la existencia de las células madre del cáncer, en este caso, en el contexto de la leucemia mieloide aguda (AML). Este

trabajo demostró, mediante el uso de un modelo murino *NOD/SCID*, la existencia de lo que ellos llamaron células iniciadoras de la leucemia mieloide aguda<sup>33</sup>.

Gracias a este descubrimiento, otros trabajos se centraron en determinar la existencia de estas CSCs en otros tipos de cáncer. Hasta el momento se han identificado CSCs en diferentes tipos de neoplasias hematológicas, incluyendo otros subtipos más específicos de AML<sup>34</sup>, así como en la leucemia linfoblástica aguda tanto de células B como de células T<sup>35,36,37</sup>. Además, también se ha podido demostrar la existencia de las CSCs en tumores sólidos, como en el caso del cáncer de mama<sup>38</sup>, glioblastoma<sup>39</sup>, cáncer de páncreas<sup>40</sup>, cáncer de estómago<sup>41</sup>, cáncer de hígado<sup>42</sup>, cáncer de cabeza y cuello<sup>43</sup>, cáncer de colón<sup>44</sup> y cáncer de próstata<sup>45</sup>.

Teniendo en cuenta esta hipótesis, la CSC es la única responsable del mantenimiento del tumor<sup>46</sup>. Así, la formación y mantenimiento del tumor vienen dados por una alteración en la diferenciación de la célula madre del cáncer y no, como se asumía, por una alteración en la proliferación celular.

Este modelo de las CSCs, también denominado “Modelo Jerárquico”, propone que los tumores malignos muestran una organización jerarquizada<sup>33</sup>, por lo que no todas las células que conforman la masa tumoral son igualmente competentes para regenerar el tumor. Dentro de la población cancerígena del tumor existen células madre del cáncer que son células tumorigénicas que además presentan unas características biológicas diferentes a las del resto de subpoblaciones del tumor<sup>47</sup>. Estas células madre del cáncer tienen dos características definitorias: su capacidad para dividirse indefinidamente y su capacidad para diferenciarse en progenie no tumorigénica que favorece el crecimiento y mantenimiento del tumor.

Existen dos hipótesis sobre el origen de las células madre del cáncer: (1) la primera hipótesis sostiene que las células madre normales sufrirían múltiples mutaciones, como resultado de un proceso de inestabilidad genética y epigenética, transformando a estas células madre normales en CSCs<sup>48</sup>; (2) la segunda hipótesis postula que las células del tumor irían adquiriendo progresivamente características de célula madre, invirtiendo su ontogenia a través de la plasticidad inducida por el oncogén<sup>49</sup>.

Desde esta perspectiva, es necesario reconsiderar nuestra forma de entender el desarrollo del cáncer, así como nuestra forma de tratarlo, dado que las terapias actuales van dirigidas a frenar la proliferación celular, no siendo eficaces frente a las CSCs. Este hecho explicaría las frecuentes recaídas de los pacientes con cáncer tras

ser tratados con las terapias convencionales<sup>50</sup>. De hecho, de la misma manera que las células madre normales continúan dando lugar a tejidos normales después de la quimioterapia (crecimiento del cabello perdido, regeneración de los compartimentos hematopoyéticos, reparación de las mucosas intestinales), las CSC también regeneran el tumor. Hay algunas excepciones en esta regla, como es el caso del carcinoma testicular, donde las células madre tumorales son más sensibles a la quimioterapia que las células madre normales, permitiendo que este tipo de cáncer se cure.

### **2.2 Implicaciones prácticas de la Teoría de la Célula Madre del Cáncer.**

Como hemos dicho, las CSCs comparten muchas características funcionales con las células madre normales, incluyendo su potencial de autorenovación ilimitado. Además, muchas de las rutas de señalización involucradas en el mantenimiento de las células madre normales están desreguladas en el cáncer (por ejemplo la vía de *Wnt*<sup>51-53</sup>, *Notch*<sup>54,55</sup>, *Shh*<sup>56</sup>, *Bmi*<sup>57,58</sup>, *Pten*<sup>59,60</sup>, etc). Esto hace que sea inevitable plantearse si las terapias dirigidas frente a las CSCs, cuya diana serían estas rutas de señalización compartida, podrían también eliminar a las células madre residentes normales<sup>4</sup>, haciendo inviables los tratamientos frente a las CSCs por su toxicidad. Sin embargo, cabe esperar que existan diferencias potenciales entre las células madre normales y las células madre del cáncer que proporcionen una ventana terapéutica<sup>61</sup>.

Estas diferencias entre las células madre normales y las CSCs han sido identificadas en diferentes modelos. Por ejemplo, en un modelo murino de leucemia se describió que las CSCs presentaban un programa transcripcional específico encargado de funciones particulares como la autorenovación<sup>62</sup>. Otro estudio en ratones demostró que el mecanismo de autorenovación para mantener el *pool* de células madre leucémicas y de células madre hematopoyéticas (HSCs) era diferente<sup>63</sup>. En este estudio, los investigadores demostraban que la delección del gen supresor de tumores *Pten* daba lugar a la generación de células madre leucémicas pero eliminaba a las células madre hematopoyéticas normales. Esta distinción mecánica abriría una puerta que permitiría diseñar terapias, para combatir la leucemia, dirigidas a la vía de *PTEN*, con el objetivo de eliminar a las células madre leucémicas sin dañar el pool de células madre hematopoyéticas normales.

Aún más relevante, ha sido la identificación, en tumores de mama humanos, de vías de señalización activadas específicas de las denominadas CSC de mama ( $CD44^+$  /  $lin^-$  /  $CD24^{-low}$ ), gracias al análisis del perfil genético de las subpoblaciones celulares de

estos tumores<sup>64</sup>. Este hallazgo permitiría explorar nuevas terapias dirigidas a inactivar estas vías de señalización específicas de las CSC de mama.

Por lo tanto, el camino a seguir es identificar, al igual que en el caso de la leucemia o el cáncer mama, aquellas características genéticas, epigenéticas, moleculares, etc, que hacen a las CSCs diferentes a las células madre normales en cada tipo de cáncer. También sería interesante encontrar qué características de las CSCs difieren de las características propias de las células del resto del tumor, con el fin de tener una visión más exacta sobre la biología de las CSCs a la hora de diseñar terapias frente a ellas<sup>61</sup>.

De este modo, la primera opción lógica que se nos podría ocurrir para estudiar la biología de las CSCs sería utilizar células madre del cáncer humano obtenidas directamente de pacientes. Este método, sin embargo, presenta importantes desventajas a la hora de identificar marcadores específicos, debido a la gran variabilidad existente entre pacientes en términos de antecedentes genéticos. Además, esta aproximación tampoco podría tener en cuenta otras variables particulares de cada individuo, como por ejemplo, la susceptibilidad al cáncer de cada paciente<sup>4</sup>. Por otro lado, se ha demostrado que los reordenamientos genómicos aparecen con una frecuencia relativamente alta en células somáticas de seres humanos sanos<sup>65</sup>, lo que introduciría una mayor variabilidad a la hora de caracterizar el comportamiento de las CSCs de cada tumor en particular.

Así, la homogeneidad de las muestras con las que se va a trabajar, a la hora de encontrar marcadores específicos de las CSCs, es un factor vital. Esto implica la necesidad de desarrollar nuevas aproximaciones prácticas, como por ejemplo el uso de modelos animales basados en la teoría de las CSCs, que nos permitan identificar con precisión las vías genéticas responsables del desarrollo del cáncer.

Una de estas aproximaciones prácticas es la generación de modelos de ratón, los cuales integran la complejidad de un órgano y sus diferentes tipos celulares dentro de la dinámica fisiológica del animal<sup>66</sup>. De este modo, reproducir la patología del cáncer humano mediante el desarrollo de modelos murinos debe ser no sólo un requisito indispensable para comprender la génesis del cáncer humano<sup>67</sup>, sino también para el desarrollo de nuevas terapias moleculares y farmacológicas que nos permitan prevenirlo y tratarlo<sup>68</sup>. De hecho, la mayor parte de nuestra conceptualización actual sobre la forma en que se produce el cáncer en los seres humanos está fuertemente influenciada por los modelos de cáncer en ratón<sup>69-72</sup>.

En la actualidad existen múltiples modelos murinos que tratan de mimetizar el cáncer humano. Uno de estos modelos son los *xenografts*, que consiste en inyectar células tumorales de origen humano en ratones inmunodeprimidos. Esta aproximación ha sido crucial para establecer y consolidar el concepto de las CSCs<sup>33,35,38,73</sup>. Sin embargo, estos modelos presentan importantes limitaciones, como por ejemplo, la necesidad de donantes humanos. Además, estos modelos no son capaces de reproducir completamente las características fundamentales de los tumores humanos, dado que los tumores desarrollados en *xenografts* no presentan interacciones con la vasculatura o el sistema inmune. Esto hace que este tipo de modelos no sean adecuados para entender las etapas iniciales del desarrollo del cáncer.

Una alternativa a estos modelos es la transfección retroviral de oncogenes humanos en células progenitoras de ratón. Esta técnica consiste en manipular genéticamente células de ratón de tejidos específicos mediante la introducción retroviral de los oncogenes humanos deseados, para a continuación, introducir estas células modificadas en huéspedes singénicos. Esta técnica permite comparar los efectos de diferentes oncogenes en un entorno celular de partida definido, y estudiar cómo un determinado oncogén puede lograr efectos distintivos en diferentes tipos celulares<sup>4</sup>. Sin embargo, esta técnica no ha aportado grandes avances en el conocimiento de las CSCs.

De manera que, los mejores modelos murinos para abordar la génesis del cáncer son aquellos en los que las modificaciones genéticas han sido introducidas en la línea germinal (ratones transgénicos o *knock-in*).

Dado que el paradigma de las CSCs es relativamente moderno, la idea de que el cáncer es una proliferación clonal de una célula diferenciada anormal que ha sufrido un cierto número de mutaciones oncogénicas y, que como resultado de las mismas, se ha expandido de una manera descontrolada, casi como una línea celular, se ha sostenido durante muchos años y aún algunos la sostienen. Esto ha hecho que se hayan generado cientos de modelos de ratones transgénicos que dirigen la expresión de los oncogenes a poblaciones celulares diferenciadas, dando lugar a síndromes hiperproliferativos o expansiones celulares que no reproducen las características del cáncer humano.

Por lo tanto, el desarrollo de modelos de ratón, basados en la teoría de CSC, induciendo la expresión de los oncogenes en progenitores tempranos, para reproducir

y entender la génesis del cáncer humano, parece indispensable<sup>18,19</sup>. El desarrollo de estos modelos nos permitiría analizar los perfiles genéticos, epigenéticos, proteómicos y moleculares de las CSCs en diferentes estadios de la enfermedad, para así compararlos con los de las células madre normales<sup>7</sup>. De esta manera, podremos obtener información muy relevante acerca de genes concretos responsables del mantenimiento de las CSCs, que nos permita desarrollar tratamientos frente a estos genes.

En resumen, la teoría de las CSCs pone de manifiesto que no todas las células que componen la masa tumoral poseen la capacidad de regenerarla, presentando una organización jerarquizada. Por lo tanto, si hoy en día se acepta la idea de que la masa tumoral proviene de las CSCs, entonces ¿de dónde proceden estas células madre del cáncer? ¿Cuál es su origen?

### **2.3 Célula de origen del cáncer vs célula madre del cáncer.**

La célula de origen del cáncer o “*Cancer cell of origin*” (CCO) en inglés, es una célula inicialmente sana (que no necesariamente tiene que ser una célula madre) que será reprogramada por el o los eventos oncogénicos y que finalmente dará lugar a una célula (pre) tumoral con propiedades de célula madre. Es importante tener en cuenta que esta célula de origen del cáncer no tiene por qué estar relacionada con la CSC (la única célula encargada de mantener el crecimiento del tumor). Por lo tanto, los conceptos de célula de origen y de célula madre del cáncer hacen referencia a la célula iniciadora del cáncer y a la célula que mantiene el cáncer, respectivamente<sup>74</sup>, por lo que no deben ser confundidos.

Desde un punto de vista evolutivo, hay dos posibilidades a considerar en este contexto de la célula de origen del cáncer. Una opción es que la célula de origen que sufre el primer evento oncogénico sea una célula madre: en este caso, una nueva célula tumoral será reprogramada para generar el nuevo tejido patológico en lugar del normal. En el caso de la CML, al restringir la expresión de la lesión oncogénica al compartimento de las células madre/progenitoras hematopoyéticas en ratones modificados genéticamente, se genera un tumor muy similar al humano, con su misma variabilidad celular<sup>18,75</sup>. En modelos de ratón de tumores intestinales, cuando la lesión oncogénica (la activación de la vía de señalización de *Wnt*) se dirige al compartimento de las células madre, se generan adenomas intestinales que mantienen su jerarquía de desarrollo; sin embargo, cuando el evento oncogénico se dirige a las células epiteliales intestinales diferenciadas, solamente aparecen pequeños microadenomas<sup>76-</sup>

<sup>78</sup>. Estos datos sugieren que el tumor debe tener su origen en las células madre de la cripta. Algo similar ocurre en los tumores del sistema nervioso; cuando las lesiones oncogénicas asociadas a los astrocitomas se dirigen a los progenitores tisulares (es decir, a la zona subventricular), los tumores se desarrollan, mientras que la astrogliosis sólo ocurre cuando las células diana son células parenquimáticas adultas diferenciadas<sup>79</sup>. Estos ejemplos y otros<sup>80,81</sup>, demuestran que el evento iniciador puede tener lugar en una célula madre normal, incluso si la masa tumoral se compone de células diferenciadas, y que las lesiones oncogénicas tienen la capacidad de reprogramar a esta célula madre, dando lugar a la aparición de una población tumoral diferenciada<sup>75</sup>.

Existe, sin embargo, otra posibilidad con respecto a la naturaleza de la célula de origen del cáncer: puede ser que una célula diferenciada sea la diana original del o de los eventos oncogénicos iniciadores, y en el posterior proceso de reprogramación tumoral mediada por los oncogenes, la célula recuperaría las características de célula madre para convertirse en una verdadera CSC. Esta posibilidad tiene, sin embargo, dos requisitos previos: por parte de la célula, aunque se trate de una célula diferenciada, debe poseer la suficiente plasticidad para ser reprogramable (al menos, por este oncogén específico)<sup>82</sup>; y por parte del oncogén, la alteración debe ser capaz de activar todos los programas necesarios para conferir a la célula propiedades de célula madre (al menos, para esta célula diana específica), en último término.

Se ha demostrado experimentalmente que algunos oncogenes son capaces de generar CSC cuando se introducen en células diana comprometidas con un linaje celular; es el caso de *MOZ-TIF2*<sup>83</sup>, *MLL-AF9*<sup>62,84</sup>, *MLL-ENL*<sup>85</sup>, *MLL-GAS*<sup>86</sup> o *PML-RaRa*<sup>87,88</sup>. *MLL-AF9*, por ejemplo, es capaz de conferir propiedades de autorenovación a los progenitores comprometidos de granulocitos-macrófagos, activando en ellos un programa de células madre<sup>62</sup>. Del mismo modo, *c-Myc* puede dar lugar a CSCs epiteliales induciendo un programa de transcripción de células madre embrionarias en células epiteliales diferenciadas<sup>89</sup>. En cambio, otros oncogenes son incapaces de conferir propiedades de autorenovación en las células sobre las que actúan, como es el caso de *BCR-ABLp190*<sup>83</sup>. En estos casos, en los que el oncogén no es capaz de conferir propiedades de célula madre, se podría originar una célula precancerosa que posteriormente pueda, con la acumulación de lesiones genéticas adicionales, adquirir propiedades de célula madre y dar lugar finalmente a una CSC<sup>90</sup>.

Desde esta perspectiva de la CSC, nuevos métodos de rastreo de linaje han tratado de descubrir el origen celular del cáncer en muchos tejidos. Uno de estos métodos es el uso de promotores específicos de tipo celular dirigidos por Cre recombinasas que han permitido la introducción de oncogenes o la eliminación de genes supresores de tumores en tejidos sanos usando modelos murinos. Estos modelos son mucho más útiles que los modelos *in vitro* o los modelos de *xenografts*, ya que presentan la organización apropiada del tejido y la presencia de los sistemas estromales, inmunes, linfáticos, nerviosos y vasculares nativos<sup>91</sup>. Este tipo de aproximación experimental ha permitido determinar la célula de origen de muchos tipos de cáncer y sigue permitiendo determinar el origen celular de otros tipos de tumores. Por ello, uno de los principales objetivos de esta tesis doctoral va a ser, por un lado, determinar la célula de origen del linfoma difuso de células B grandes (DLBCL) asociado a la pérdida del gen *CREBBP*<sup>92</sup> y, por otro lado, determinar el origen celular de la leucemia linfoblástica aguda de células T (T-ALL) *Lmo2+*, de la que hasta ahora mayoritariamente se sostiene que es una célula progenitora de linaje T.

### 3. El mecanismo de reprogramación tumoral como origen del cáncer.

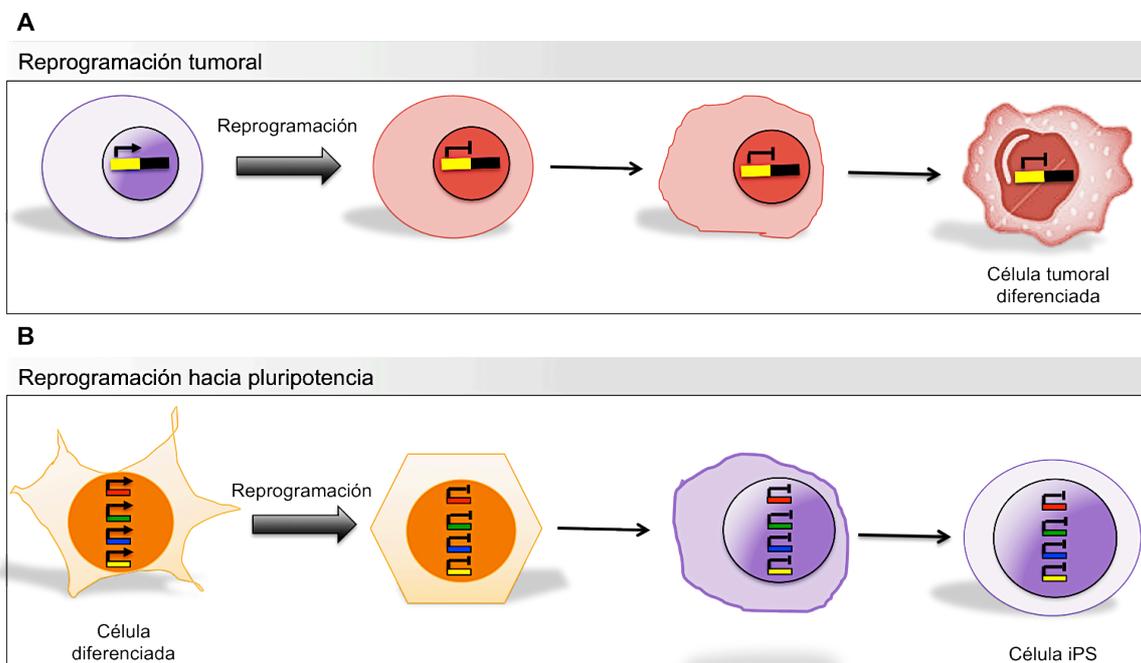
#### 3.1 La reprogramación como origen del cáncer.

Tal y como hemos mencionado en los apartados anteriores, el cáncer es una enfermedad heterogénea, lo que hace pensar que probablemente el comportamiento de los oncogenes dentro de este contexto también pueda ser heterogéneo, dependiendo de la célula diana en la que actúen. Esta observación es compatible con la teoría de la célula madre del cáncer, por lo que podríamos considerar el cáncer como una enfermedad que surge y se mantiene del mismo modo que lo hace el sistema hematopoyético o cualquier otro tejido dirigido por células madre: el programa genético de las células madre es todo lo necesario para constituir o reconstituir toda la heterogeneidad celular del tejido, y la información responsable de este programa no necesita estar presente en las células diferenciadas que forman el tejido<sup>93</sup>.

La reprogramación celular se refiere a “el concepto de re-escribir la red epigenética y transcripcional que define el estado de una célula a la de un tipo de célula diferente”<sup>94</sup>. Cuando este proceso es inducido por medios molecularmente definidos, se denomina reprogramación directa. El término reprogramación se utiliza con bastante frecuencia, ya que es el mecanismo encargado de dirigir la generación de células madre

pluripotentes inducidas (iPSCs) por medio de un cóctel específico de factores de transcripción<sup>95,96</sup>.

Sin embargo, como se ha definido anteriormente, el concepto de reprogramación también abarca otros procesos de cambio de destino de la célula, que no necesariamente implican a los cuatro factores de Yamanaka, ni a la generación de iPSCs como destino final. Por consiguiente, la reprogramación tumoral es el proceso por el cual un oncogén o una alteración genética puede restablecer el estado epigenético y/o el transcriptómico de una célula inicialmente sana (célula de origen del cáncer), estableciendo, así, un nuevo programa de diferenciación, patológico, que finalmente conduce al desarrollo del cáncer (**Figura 3**).



**Figura 3: Reprogramación tumoral vs reprogramación hacia pluripotencia. A)** Recientes evidencias genéticas *in vivo* han mostrado que los oncogenes humanos son capaces de reprogramar células madre y/o progenitores tempranos hacia identidades específicas de células diferenciadas tumorales, pero su presencia no es requerida después, dentro de la masa tumoral. **B)** El proceso de reprogramación también aparece en otros contextos fuera del cáncer, como durante la formación de células madre pluripotentes inducidas (iPSCs) *in vitro*. Sin embargo, a diferencia de la reprogramación de las células madre tumorales, el proceso de iPSCs se inicia en una célula diferenciada.

Teniendo en cuenta esta hipótesis, el cáncer puede tener su origen en la reprogramación tumoral de una célula madre. Así, el oncogén o alteración genética sería responsable de la reprogramación de la célula de origen del cáncer, pero una vez reprogramada dicha célula, la alteración genética no sería necesaria para el

mantenimiento del tumor. Esto introduce una nueva visión a la hora de entender la transformación oncogénica inicial, como un proceso pasajero que no tendría por qué llevar a cabo una función en la expansión y mantenimiento del tumor<sup>7,97</sup>.

Por lo tanto, este mecanismo de reprogramación tumoral guarda un gran paralelismo con el proceso de pluripotencia<sup>18,98</sup>, ya que en ambos procesos se requiere la expresión o supresión transitoria de genes específicos (*c-Myc*, *Sox2*, *KLF4*, etc)<sup>95</sup>. Además, muchos de los factores encargados de inducir pluripotencia están asociados con la oncogénesis<sup>99-102</sup>.

Esta relación entre la reprogramación tumoral y el proceso de pluripotencia también pone en evidencia que existen genes que tienen una actividad lo suficientemente fuerte para modificar el mecanismo celular encargado de mantener la identidad de la célula y forzar un cambio en ésta. Sin embargo, la mayoría de genes carecen de esta fuerte actividad y sólo unos pocos reguladores epigenéticos o transcripcionales tienen la capacidad de reprogramar una célula<sup>103</sup>.

Un aspecto conceptual y biológico importante de los mecanismos moleculares de reprogramación hacia la pluripotencia es el hecho de que, una vez que se ha iniciado el camino a la pluripotencia y se ha producido la reprogramación *per se*, la condición pluripotente ya no requiere de factores inductores para su mantenimiento<sup>104</sup>. Así, por paralelismo, si las CSCs son generadas por un proceso de reprogramación (tumoral), entonces tal vez los oncogenes que inician la formación del tumor pueden no ser necesarios para la progresión del mismo<sup>18,98</sup>. Teniendo en cuenta esta hipótesis, la lesión iniciadora sería la fuerza impulsora en el proceso de reprogramación, esencial para la tumorigénesis. Sin embargo, una vez que ha tenido lugar la reprogramación, este evento iniciador sólo sería una mutación pasajera dentro de las CSCs, ya sea sin una función significativa, o desempeñando un papel diferente, sin relación con la reprogramación, en la expansión o proliferación tumoral. Este modo de acción explicaría por qué las terapias dirigidas, bien diseñadas, fracasan en la erradicación de la enfermedad, a pesar de su aparente eficacia frente a la masa tumoral principal<sup>7</sup>.

Por lo tanto, a la hora de entender el cáncer y tratarlo, debemos tener en cuenta tres aspectos fundamentales: la célula de origen, el proceso de reprogramación tumoral y la célula que mantiene el tumor (CSC). Sin embargo, para llegar a tener un conocimiento pleno de estos tres aspectos clave, es necesario estudiar si otros factores pueden estar influyendo en su comportamiento.

### **3.2 Modelos experimentales in vivo de reprogramación tumoral de células madre.**

Para poder demostrar la falta de homogeneidad en el comportamiento de los oncogenes en la biología del tumor, sería necesario diseccionar y aislar la función que el oncogén está jugando en las primeras etapas de la enfermedad, a nivel de la célula de origen. De hecho, para demostrar que el mantenimiento de la expresión del oncogén no es necesario para la progresión tumoral, más allá del paso inicial de la reprogramación, habría que encontrar una manera de restringir la expresión del oncogén al compartimento progenitor. Este sistema también nos permitiría probar, si éste fuera realmente el caso, que los oncogenes que inician la formación del tumor podrían ser dispensables para la progresión y/o mantenimiento tumoral.

Existen diferentes evidencias genéticas *in vivo* que demuestran que el cáncer es un tejido dirigido por células madre. Para demostrarlo, en estos trabajos, utilizan la región control del locus del gen de ratón *Sca1* (*Ly6E.1*) para restringir la expresión de oncogenes seleccionados, específicos del cáncer humano, al compartimento de células madre hematopoyéticas en un ratón transgénico<sup>18-23</sup>. En un principio, estos estudios se centraron en el gen de fusión *BCR-ABLp210* asociado a la CML, debido a que la comunidad científica acepta que este tipo de cáncer es un trastorno de las células madre hematopoyéticas. La CML comienza con una fase crónica prolongada, que tras una media de 5 años (en ausencia de tratamiento), se convierte en una crisis blástica, indistinguible de una leucemia aguda. El tratamiento que actualmente se usa en la práctica clínica es un inhibidor específico de BCR-ABL, STI571 (*imatinib*), que es capaz de eliminar la masa tumoral de células diferenciadas que expresan *BCR-ABL*, pero es incapaz de eliminar la pequeña porción de CSC que expresan *BCR-ABL* y mantienen el tumor.

En ratones transgénicos, cuando la expresión de *BCR-ABLp210* está restringida al compartimento de células madre/progenitoras hematopoyéticas, los ratones *Sca1-BCR-ABLp210* desarrollan una CML que recapitula las principales características de la enfermedad humana. En estos ratones, según el diseño experimental, aunque la iniciación tumoral tiene lugar en el compartimento de célula madre/progenitora, todas las células diferenciadas leucémicas, que forman la masa principal del tumor, han apagado la expresión del oncogén. Por lo tanto, *BCR-ABL* no es expresado en las células hematopoyéticas diferenciadas, ni siquiera en las tumorales<sup>18</sup>. Con estos datos, la duda lógica que surge es ¿cómo es posible que la CML desarrollada por los ratones *Sca1-BCR-ABLp210* se produzca de una manera diferente, cuando en las CML

humanas todas las células leucémicas (no sólo las CSC) portan las lesiones oncogénicas? Esto sólo puede explicarse si *BCR-ABL* está reprogramando realmente, mediante un mecanismo “*hit-and-run*”, a las células madre/progenitoras hematopoyéticas para su malignidad. En resumen, estos resultados representaron la primera prueba genética *in vivo* que asoció mecanísticamente la tumorigénesis y la reprogramación de células madre/progenitores tempranos, y descubrieron la posibilidad de un mecanismo similar a la reprogramación en el desarrollo del cáncer. Otros ejemplos de reprogramación tumoral de células madre, han sido descritos para otros tipos de tumores. Por ejemplo, el gen de fusión *EWS-FLI-1*, asociado con el 85-90% de los tumores de la familia del sarcoma de Ewing, induce la expresión de los genes *OCT4*, *SOX2* y *NANOG*, asociados a células madre embrionarias, en células madre mesenquimales pediátricas humanas pero no en sus homólogos adultos, reprogramándolas hacia CSC de sarcoma de Ewing<sup>105</sup>. Otro ejemplo es el del oncogén *SYT-SSX2* asociado al sarcoma sinovial, que es capaz de reprogramar las células mesenquimales dictando su compromiso con un linaje pro-neural, haciendo que probablemente esta lesión genética constituya el evento iniciador en este tipo de tumor<sup>106</sup>.

#### **4. CREBBP y su función en el desarrollo del linfoma folicular y del linfoma difuso de células B grandes.**

##### **4.1 Los linfomas no Hodgkin: definición, epidemiología, clínica y clasificación.**

Los linfomas no Hodgkin (NHL), también denominados como linfomas malignos para diferenciarlos del linfoma de Hodgkin (cuyas características y pronóstico difieren profundamente), son neoplasias hematológicas, clásicamente definidas por tener su origen en células linfoides (que pueden ser células B, T o linfocitos *natural killer*) en diferentes estadios de maduración. La mayoría de los linfomas no Hodgkin (alrededor del 85%) se piensa que tienen su origen en una célula B madura. En los NHL, una célula linfoide con un determinado estadio madurativo queda detenida, reproduciéndose de forma descontrolada y causando al cabo del tiempo la hipertrofia del órgano en el que se produce. Dado que el tejido linfático se extiende por todo el cuerpo, los linfomas pueden surgir en cualquier parte del mismo, y a partir de aquí, diseminarse por otros órganos y tejidos. En la mayoría de los casos el linfoma comienza con la infiltración de un ganglio linfático (son las denominadas formas

nodales), pero algunos subtipos específicos pueden surgir en la piel, cerebro, bazo u otros órganos (formas extranodales).

Los linfomas no hodgkinianos suponen el 9,8% de todos los tipos de cáncer en Europa. En nuestro país aparecen entre 30-70 nuevos casos por millón de habitantes y año, presentando una incidencia similar a la media europea del 10% (*datos obtenidos de eco.iarc.fr/eucan*). Son algo más frecuentes entre los hombres y entre pacientes con enfermedades del sistema inmune (SIDA, inmunodeficiencias, receptores de trasplantes de órganos, enfermedades autoinmunes), infecciones (gastritis por *Helicobacter*, virus de Epstein Barr), y pacientes tratados con quimioterapia o radioterapia. Aunque de forma mucho menos frecuente, es un tipo de cáncer que también puede afectar a los niños.

Los síntomas que presentan los pacientes con linfoma no Hodgkin son diversos y varían en función del lugar de afectación de la enfermedad, el tamaño del tumor y la velocidad de crecimiento del mismo. En muchas ocasiones los pacientes presentan los denominados síntomas B, que incluyen fiebre, sudores nocturnos y pérdida de peso. Además, la enfermedad también suele cursar con fatiga, molestias en el pecho y dificultad para respirar debido a la aparición de adenopatías en el mediastino.

El término de linfoma no Hodgkin engloba más de 20 tipos distintos de cáncer<sup>107</sup> con características individuales, pero que pueden clasificarse en dos grandes grupos en función de su agresividad:

- Los **linfomas de alto grado o agresivos**, que tienden a crecer y extenderse rápidamente y provocan síntomas graves. Los más frecuentes son: linfoma difuso de células B grandes ó DLBCL (es el tipo de linfoma más frecuente y supone entre un 30-40% de todos los linfomas<sup>108</sup>), linfoma de células del manto, linfoma T periférico, linfoma de Burkitt, linfoma linfoblástico (equivalente a una leucemia linfoblástica aguda con nula o escasa presencia de linfoblastos en sangre).
- Los **linfomas de bajo grado o indolentes**, con un comportamiento menos agresivo, se caracterizan por el desarrollo de adenopatías que pueden tener años de evolución y un estado general conservado, a pesar de hallarse, por lo general, muy extendidos. Los más frecuentes son: linfoma folicular (siendo el segundo tipo de linfoma más frecuente,

por detrás del DLBCL), linfoma de linfocitos pequeños (equivalente a una leucemia linfocítica crónica ó CLL), linfoma linfoplasmocítico o macroglobulinemia de Waldenström, linfoma de la zona marginal (en el que se incluyen los linfomas MALT) y los linfomas cutáneos de células T (micosis fungoide y síndrome de Sézary).

Paradójicamente, los linfomas no Hodgkin de alto grado suelen responder bien a los tratamientos quimioterápicos, a diferencia de los linfomas indolentes que son muy difíciles de erradicar por completo.

Mientras que para la mayoría de los cánceres la incidencia y la mortalidad han ido disminuyendo a lo largo de las últimas décadas, en el caso del linfoma no Hodgkin estos dos factores han aumentando constantemente a lo largo del tiempo<sup>109</sup>. Esto ha puesto de manifiesto la necesidad de desarrollar nuevas aproximaciones que nos permitan entender mejor la biología de esta enfermedad, poniendo especial hincapié en los dos tipos de linfoma no Hodgkin más frecuentes: el linfoma difuso de células B grandes y el linfoma folicular.

#### ***4.1.1 El linfoma folicular: definición, epidemiología, inmunofenotipo, clínica, y etiología.***

El linfoma folicular (FL) es el tipo de linfoma no Hodgkin de bajo grado más frecuente, presentando una incidencia alrededor del 20% dentro de todos los casos de linfoma no Hodgkin. Se trata de un linfoma de células B, en el que las células tumorales crecen en grupos para formar nódulos. Morfológicamente, el FL se define como una proliferación de células B malignas del centro germinal (GCB). Puede aparecer a cualquier edad, aunque es propio de personas adultas, siendo la edad media en el diagnóstico 60 años<sup>110</sup>. Es un tipo de linfoma muy poco frecuente en jóvenes y niños.

Habitualmente, los pacientes con FL no presentan síntomas y suelen acudir a su médico tras percibir una adenopatía indolora, lo que hace difícil su diagnóstico temprano. El linfoma folicular se caracteriza por presentar el siguiente inmunofenotipo, que permite su diagnóstico: CD20/CD79a+, Bcl2+, Bcl6+, CD10+. El 60% de los pacientes presenta infiltración de la médula ósea al diagnóstico. Además, alrededor del 25% de los enfermos de FL presenta esplenomegalia y una tercera parte puede tener infiltración en el hígado. La evolución de este tipo de linfoma es indolente y lenta, presentando una supervivencia media que supera los 12 años desde el diagnóstico. Lamentablemente, la curación de estos pacientes es difícil y para los enfermos en un estadio avanzado (que suponen el 80% de los casos) es imposible alcanzarla con los

tratamientos convencionales. Aunque una parte de los pacientes puede alcanzar la remisión completa, a medida que avanza el curso del FL, las recaídas son más frecuentes y los periodos en remisión más cortos. En el 25-35% de los pacientes enfermos de linfoma folicular, la enfermedad se transforma en un linfoma agresivo<sup>111,112</sup> como consecuencia de alteraciones moleculares, dando lugar con frecuencia a otro tipo de linfoma denominado linfoma difuso de células B grandes (DLBCL).

Se piensa que el FL surge en los centros germinales de los tejidos linfoides, donde las células B sufren una expansión clonal y múltiples modificaciones genéticas a través de los procesos de hipermutación somática y de cambio de isotipo<sup>113</sup>. A nivel molecular, entre el 85-90% de los linfomas foliculares se caracterizan por presentar la translocación t(14;18)(q21;q32), la cual pone al proto-oncogén antiapoptótico *BCL2* bajo el control transcripcional del locus de la inmunoglobulina pesada (*IGH*), conduciendo finalmente a la sobreexpresión de *BCL2*<sup>114-117</sup>. Sin embargo, la t(14;18) por sí sola no es suficiente para producir la enfermedad, tal y como se ha podido comprobar en individuos sanos con dicha translocación, los cuales nunca desarrollaban linfoma folicular<sup>118,119</sup>. El hecho de que la t(14;18) sea necesaria pero no suficiente para inducir FL también ha sido apoyado por los experimentos *in vivo*, en los que el linfoma folicular no se desarrollaba al inducir la translocación t(14;18) en células B de ratón<sup>120,121</sup>. Todo ello ha provocado que la mayoría de los esfuerzos en el estudio del linfoma folicular se hayan centrado en identificar eventos genéticos adicionales que cooperen con la t(14;18) para inducir la transformación maligna. Además, el 10-15% de casos de FL de tipo adulto restantes no albergan la translocación t(14;18)<sup>122</sup>, y el mecanismo patogénico que dirige el desarrollo de este subgrupo de casos no se conoce suficientemente bien.

Los estudios del genoma<sup>123</sup>, los análisis citogenéticos<sup>124,125</sup>, los estudios de expresión génica<sup>126,127</sup> y los perfiles de microARNs<sup>128,129</sup> nos han permitido tener un mejor conocimiento de la genética del linfoma folicular. Sin embargo, el cambio más significativo en nuestra comprensión de la genética del linfoma folicular se ha producido a partir de los avances en las técnicas de secuenciación masiva (NGS), que nos han permitido relacionar la enfermedad con mutaciones en un sorprendente número de reguladores epigenéticos como *CREBBP*, *EP300*, *EZH2*, *KMT2D*, *MEF2B*, varios miembros del complejo de remodelación del nucleosoma (SWI/SNF), incluyendo *ARID1A*, *ARID1B* y *BCL7A*, así como los miembros de las familias de genes enlazadores de la histona H1 y de la histona H2<sup>130-138</sup>. Con el tiempo, la

desregulación epigenética ha sido reconocida como una característica crucial de muchos cánceres sólidos y hematológicos<sup>139</sup>. Sin embargo, lo que distingue al FL es su aparente predisposición a adquirir alteraciones en reguladores epigenéticos que además aparecen en la mayoría de los pacientes con este tipo de linfoma<sup>135,140</sup>.

A pesar de que las alteraciones epigenéticas son comunes en el linfoma folicular de individuos adultos, esta característica no suele aparecer en el linfoma folicular ganglionar de tipo pediátrico (PTNFL), una variante rara de linfoma folicular t(14; 18)-negativo que presenta mutaciones recurrentes en *TNFRSF14*<sup>141</sup> y en genes de la ruta de señalización de las MAPK<sup>142</sup>. La aparición poco frecuente de alteraciones epigenéticas, junto con un pronóstico mucho más favorable de los PTNFL comparados con los FL adultos, apoyaría la idea de que la inestabilidad epigenética que encontramos en los linfomas foliculares adultos podría ser la causa de la durabilidad y del curso clínico de la enfermedad.

Todo ello ha contribuido a que muchos de los estudios para entender la etiología de la enfermedad se hayan centrado en buscar mutaciones en reguladores epigenéticos como “*drivers*” de la enfermedad. Uno de los “*drivers*” epigenéticos que más frecuentemente se encuentra mutado en el linfoma folicular es el gen *CREBBP*<sup>132,143</sup>. Además, mutaciones en dicho gen han sido halladas en pacientes con DLBCL<sup>132</sup>, lo que sugiere que ambas enfermedades comparten eventos patogénicos.

#### **4.1.2 El linfoma difuso de células B grandes: definición, epidemiología, clasificación, inmunofenotipo, clínica, y etiología.**

El linfoma difuso de células B grandes (DLBCL) engloba un grupo de alteraciones hematológicas de tipo linfoide que se caracterizan por presentar células grandes con núcleos vesiculares, nucléolos prominentes, un citoplasma basófilo y una tasa de proliferación generalmente alta<sup>144</sup>. El DLBCL puede desarrollarse tanto en regiones ganglionares como en regiones extra-ganglionares.

Es la neoplasia linfoide más común en adultos. Se trata de un linfoma no hodgkiniano de alto grado o agresivo, con una prevalencia alrededor del 30% de todos los NHL y del 37% de todos los tumores de células B. Actualmente, en Europa y Estados Unidos, su incidencia es de 15-20 nuevos casos por cada 100.000 habitantes/año<sup>145</sup> y aumenta con la edad. Aunque se observa a cualquier edad, la edad media de los pacientes con linfomas de células B grandes es de 55 años.

A la hora de clasificar los DLBCL debe hacerse una distinción entre los casos derivados *de novo* (referidos como primarios) y los que representan la progresión o transformación de un linfoma menos agresivo (referidos como secundarios), por ejemplo de un linfoma folicular, de un linfoma de la zona marginal, etc. En la actualidad la OMS ha subdividido los DLBCL en cuatro categorías, algunas de las cuales no aparecían en las clasificaciones anteriores:

- a) DLBCL no especificado (DLBCL-NOS): es la forma más común de DLBCL y también de NHL y es la categoría donde centraremos parte del trabajo de esta tesis doctoral debido a su prevalencia y heterogeneidad.
- b) DLBCL con ubicación extraganglionar predominante.
- c) Linfoma de células B grandes terminalmente diferenciadas.
- d) Casos límite.

Cada categoría incluye, además, subtipos morfológicos y/o clínico-patológicos que hacen que la organización de estas neoplasias sea bastante compleja<sup>144</sup>.

El inmunofenotipo de los DLBCL-NOS es  $slg^{+/-}$  (inmunoglobulina de superficie),  $clg^{-/+}$  (inmunoglobulina citoplasmática) y expresan antígenos asociados a células B como CD19, CD20, CD22 y CD79a<sup>146</sup>. Este tipo de linfomas, además, suelen ser CD45<sup>+</sup> y PAX5<sup>+</sup>. Otros marcadores como CD10, BCL6, IRF4, LMO2, GCET1, y FOXP1 se expresan de forma variable en el DLBCL-NOS y su combinación ha supuesto la base para la construcción de algoritmos, con el objetivo de generar perfiles expresión génica que nos permitan clasificar de una manera mucho más específica los subgrupos de DLBCL-NOS<sup>147-149</sup>. Así, estos algoritmos han permitido diferenciar dos tipos moleculares distintos de la enfermedad: los *germinal center B-cell like* DLBCL o GCB-DLBCL, que se caracterizan por expresar genes asociados a las células B del centro germinal; y los *activated B-cell-like* DLBCL o ABC-DLBCL, caracterizados por la expresión de genes asociados a la activación de células B periféricas y la mitogénesis<sup>150</sup>.

Los síntomas de los DLBCL son muy variables y dependen de cada tipo específico de linfoma. Un gran porcentaje de pacientes son diagnosticados al detectarse una adenopatía. Normalmente, los pacientes sintomáticos pueden presentar fiebre, sudoración nocturna, pérdida de peso, fatiga e infecciones de repetición. También pueden producirse manifestaciones como consecuencia del crecimiento del tamaño del bazo (molestias abdominales), de la compresión de un órgano por un tumor de

gran tamaño (tos, dolor lumbar o abdominal) o del mal funcionamiento de un órgano como consecuencia de la infiltración de células cancerosas.

Los factores que predisponen al desarrollo del DLBCL incluyen agentes que producen aberraciones moleculares, así como estados de inmunodeficiencia congénita y adquirida.

Entre las principales alteraciones genéticas que aparecen en los DLBCL-NOS encontramos que en el 30-40% de los casos se detectan translocaciones en la región 3q27 que dan lugar al reordenamiento del gen *BCL6*, un represor transcripcional que es expresado de forma selectiva por las células B del centro germinal (GC), controlando la formación del mismo<sup>151-153</sup>. Dicha translocación inhibe la *down*-regulación de *BCL-6* requerida para promover la diferenciación de las células B del GC, y crea un error en el ADN que da lugar a la inactivación funcional de *p53*<sup>154,155</sup>. Además de por translocaciones, el gen *BCL6* puede aparecer alterado en los DLBCL como consecuencia de mutaciones somáticas en la región 5' no codificante del gen.

Otra de las alteraciones genéticas que más frecuentemente aparecen en el DLBCL-NOS es la translocación del gen *BCL2*, alteración característica del linfoma folicular. Esta translocación aparece en el 20% de los casos de DLBCL surgidos *de novo* y en la mayoría de los casos de DLBCL que surgen como consecuencia de la transformación de un linfoma folicular<sup>156</sup>.

En torno al 10% de los pacientes con DLBCL presentan reordenamientos del gen *C-MYC*. Aproximadamente el 20% de los casos con alteraciones en el gen *MYC* presentan, además, la translocación de *BCL2* y/o alteraciones en *BCL6* o ambas<sup>157</sup>, lo que se denomina “doble hit” o “triple hit”<sup>158</sup>.

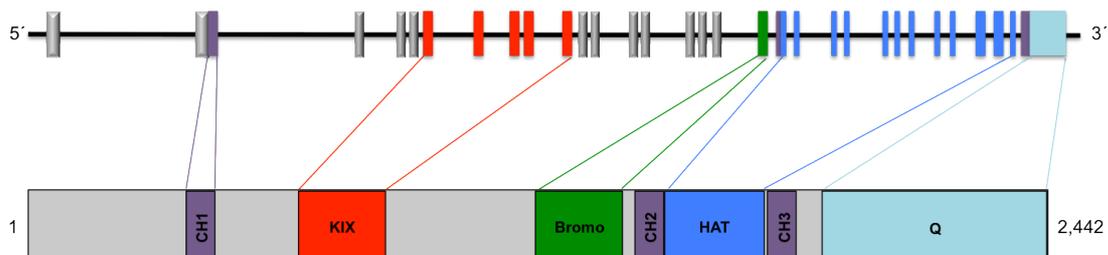
Sin embargo, el DLBCL es una enfermedad muy heterogénea, también desde el punto de vista molecular, de modo que estudios recientes han identificado en pacientes con la enfermedad, mutaciones en genes modificadores de la cromatina, tal y como ocurre en los pacientes con FL<sup>131,159,160</sup>. Entre estos genes se ha podido observar que cerca del 39% de los GCB-DLBCL y en torno al 17% de los ABC-DLBCL presentan deleciones monoalélicas y mutaciones inactivantes del gen *CREBBP*<sup>132</sup>. Este tipo de mutaciones en el gen *CREBBP* también aparece en los FL, tal y como se explicó en el apartado anterior, lo que pone aún más de manifiesto que ambas enfermedades presentan eventos patogénicos comunes<sup>132</sup>.

Por lo tanto, el linfoma difuso de células B grandes es una enfermedad agresiva y muy heterogénea desde un punto de vista genético, fenotípico y clínico. El estudio del origen de esta heterogeneidad representa un paso fundamental para nuevos avances en la comprensión de esta enfermedad, que sigue siendo un reto clínico en aproximadamente un tercio de los pacientes.

## 4.2 La proteína CREBBP humana: descubrimiento, estructura, función y patogénesis.

La proteína de unión CREB, también conocida como CREBBP o CBP, fue descubierta por primera vez debido a su capacidad para interactuar con proteínas de unión a elementos de respuesta al adenosín monofosfato cíclico ó AMPc<sup>161,162</sup>. Esta proteína está codificada por el gen *CREBBP* que se expresa ubicuamente y está implicado en la coactivación transcripcional de diversos factores de transcripción. Este gen es conocido por jugar un papel crítico en el desarrollo embrionario, el control de la proliferación celular y la homeostasis, por medio de la reorganización de la cromatina, para el reconocimiento de los factores de transcripción.

El gen *CREBBP* se localiza en una región de 156 kb situada en el cromosoma 16p13.3. Este gen está constituido por 31 exones que se transcriben en un ARNm codificante de una longitud de 2442 aminoácidos. El gen *CREBBP*, como dijimos anteriormente, codifica para la proteína de unión CREB, la cual es una proteína nuclear. Esta proteína posee una actividad intrínseca histona-acetiltransferasa (HAT) y actúa a su vez como “andamio” para estabilizar interacciones de otras proteínas con el complejo transcripcional. Esta proteína acetila tanto a proteínas histonas como a proteínas no histonas, presentando múltiples dominios de unión a proteínas (**Figura 4**): tres regiones ricas en cisteína-histidina (CH1-CH2-CH3), el dominio de interacción con CREB (KIX), el dominio de unión acetil-lisina o bromodominio (Bromo) y el dominio con actividad histona acetil-transferasa (HAT). Por último, esta proteína presenta al final de su estructura un dominio de poliglutamina (Q).



**Figura 4: Estructura de la proteína CREBBP. Esquema del gen *CREBBP* (parte superior) y de la proteína CREBBP (parte inferior).** El código de color de los exones corresponde con el código de color del dominio funcional de la proteína: dominio cisteína-histidina (CH); dominio de unión a CREB (KIX); bromodominio (Bromo); dominio histona acetil-transferasa (HAT); región de poliglutamina (Q).

La actividad acetil-transferasa de *CREBBP* está regulada de una manera dependiente del ciclo celular, de modo que, el incremento de la actividad acetil-transferasa que *CREBBP* experimenta en el límite entre la fase G1 y la fase S del ciclo celular es probablemente consecuencia de la fosforilación de CREBBP mediada por la vía de las ciclinas/Cdk (quinasas dependientes de ciclinas) antes del inicio de la fase S<sup>163</sup>.

La proteína CREBBP presenta una gran homología con otra proteína con actividad histona-acetiltransferasa, la proteína p300, lo que las hace ser parcialmente redundantes. Sin embargo, la proteína p300 fue aislada por primera vez por su interacción con la proteína adenoviral E1A<sup>164</sup>. Así, *CREBBP* y *p300* son coactivadores de al menos 40 factores de transcripción diferentes<sup>165</sup>. El gran número de proteínas que interactúan con CREBBP pone de manifiesto la influencia de esta proteína en las funciones celulares esenciales. CREBBP regula varios procesos biológicos fundamentales, incluyendo la proliferación y ciclo celular<sup>165</sup>, la diferenciación celular<sup>166,167</sup> y la respuesta al daño del ADN<sup>168</sup>. Pero la versatilidad de esta proteína también ha dificultado el conocimiento de sus implicaciones específicas en distintos procesos biológicos y estados fisiopatológicos<sup>169</sup>.

Las funciones de CREBBP en la regulación de la progresión del ciclo celular están mediadas, en parte, por su influencia en la transcripción. Sin embargo, CREBBP también puede regular el ciclo celular a través de interacciones o la acetilación de proteínas implicadas en la progresión del ciclo celular, tales como la maquinaria de replicación del ADN y las histonas. CREBBP puede regular la replicación del ADN mediante la modificación de las histonas que rodean los orígenes de replicación. Gracias a su actividad acetil-transferasa, CREBBP es responsable de la acetilación de la histona H3 a nivel de la lisina 18 (H3K18ac) y a nivel de la lisina 27 (H3K27ac), y las modificaciones a las que da lugar están asociadas con promotores activos y *enhancers*<sup>170,171</sup>. Por lo tanto, CREBBP regula varios aspectos del proceso de replicación del ADN, incluyendo la elección y el momento del inicio de replicación, así como el ensamblaje del ADN recién sintetizado en cromatina. Además, CREBBP participa en otras fases del ciclo celular, por ejemplo, se ha demostrado que la pérdida de *CREBBP* lleva a un retraso en la mitosis y a una acumulación de células en fase G2/M debido a que CREBBP es requerido para la función del complejo promotor de la anafase (APC/C), una ubiquitina ligasa E3 requerida para la progresión de la célula a través de la mitosis<sup>172</sup>. Todo ello sitúa a CREBBP en múltiples fases a lo largo del ciclo celular y pone de manifiesto las funciones de esta proteína en la promoción de la progresión a través del ciclo celular<sup>169</sup>.

En consonancia con el papel crítico de CREBBP en la regulación del ciclo celular, defectos de esta proteína dan lugar a importantes alteraciones en el crecimiento. Así, el modelo de ratón nulo para *CREBBP* es letal para el embrión<sup>173</sup>.

Aunque algunos estudios han implicado al gen *CREBBP* en el desarrollo del cáncer, la contribución de esta histona acetil-transferasa en el fenotipo del cáncer se ha definido con poca precisión. Esto se debe, en parte, a que *CREBBP* participa en diversas y, a veces, antagónicas vías celulares, tales como los procesos de supresión tumoral y los procesos oncogénicos. Además, la dificultad aumenta cuando se trata de comprender las consecuencias de las mutaciones en este gen, dado que se han identificado en múltiples estudios sobre el genoma del cáncer. Para entender completamente cómo las alteraciones genéticas y/o epigenéticas de *CREBBP* contribuyen al fenotipo del cáncer, es importante determinar qué vías celulares están afectadas específicamente como consecuencia de dichas alteraciones.

Los primeros indicios del papel como supresor tumoral de *CREBBP* provienen de los hallazgos genéticos encontrados en un desorden congénito raro del desarrollo, el síndrome de *Rubinstein-Taybi* (SRT)<sup>174</sup>. Mutaciones germinales heterocigotas de *CREBBP* se observan en el SRT y parecen desempeñar un papel importante en la patogénesis de la enfermedad. Los pacientes con el síndrome de *Rubinstein-Taybi* tienen una mayor incidencia de cáncer, lo que hace que se considere un síndrome de susceptibilidad a cáncer<sup>175</sup>. Las mutaciones de *CREBBP* en el SRT son variables y abarcan microdeleciones, mutaciones sin sentido o *nonsense*, así como mutaciones puntuales en diferentes dominios. Algunas de estas lesiones genéticas reducen la actividad acetiltransferasa y/o transcripcional de *CREBBP*, haciendo que la reducción de estas funciones esté implicada en la etiología de las neoplasias malignas asociadas con el síndrome de *Rubinstein-Taybi*<sup>176</sup>. Esta afirmación, además, ha sido apoyada por los estudios del gen *CREBBP* en ratón<sup>177-179</sup>.

Sin embargo, múltiples análisis de muestras humanas han puesto de manifiesto la importancia de las mutaciones somáticas del gen *CREBBP* en el desarrollo de algunos tipos de cáncer (**Tabla I**). Un estudio de 134 pacientes con linfoma no *Hodgkin* de células B (concretamente de linfoma difuso de células B grandes) identificó mutaciones heterocigotas de pérdida de función del gen *CREBBP* en el 29% de las muestras analizadas<sup>132</sup>. En otro estudio, en el que se analizaron 154 muestras de cáncer de pulmón humano, se identificaron mutaciones en el gen *CREBBP* en el 5-10% de las muestras<sup>180</sup>. Estas mutaciones de *CREBBP* asociadas al cáncer afectaban

a varios exones del gen, sin que hubiera una conexión evidente entre el dominio alterado de la proteína y los fenotipos observados en cada caso. De modo que, estos estudios sugieren que las mutaciones de *CREBBP* pueden tener un papel en el desarrollo de estos tumores. Esta idea se ha visto reforzada por estudios en leucemia linfoblástica aguda (ALL) donde se tomaron células de 71 pacientes al momento del diagnóstico y de nuevo en el momento de la recaída, y se detectó que aquellos pacientes que adquirirían resistencia al tratamiento presentaban mutaciones de pérdida de función en el gen *CREBBP*<sup>181</sup>. Todos estos estudios han contribuido a fortalecer la idea de que *CREBBP* funciona como un supresor tumoral.

**Tabla I: Tipos de cáncer con elevada frecuencia de mutaciones del gen *CREBBP*.**

Tipo de cáncer	N	Muestras mutadas (%)
<b>Linfoma folicular</b>	66	33,3
<b>Carcinoma de células escamosas de piel</b>	77	28,6
<b>Linfoma de células B marginales</b>	15	13,3
<b>Linfoma difuso de células B grandes</b>	242	12
<b>Carcinoma de la glándula salivar</b>	63	9,5
<b>Carcinoma de vejiga</b>	438	8,9
<b>Cáncer de endometrio</b>	337	8,0
<b>Carcinoma de pulmón microcítico</b>	52	7,7
<b>Cáncer de mama (ER+)</b>	80	7,5

*Datos obtenidos de la base de datos pública de COSMIC (cancer.sanger.uk)*

En este sentido, algunos estudios han demostrado que *CREBBP* es el segundo gen modificador de la cromatina más frecuentemente mutado en los dos tipos de linfoma no *Hodgkin* más comunes, el linfoma folicular (FL) y el linfoma difuso de células B grandes (DLBCL)<sup>131,143,182-184</sup>, por detrás de *KMT2DK*, y suele ir asociado con alteraciones en el gen antiapoptótico *BCL2*<sup>131</sup>. Algunos estudios sobre el linfoma folicular han mostrado, además, que las mutaciones de *CREBBP* ocurren de forma temprana durante la evolución clonal de la enfermedad, y dichas mutaciones se mantienen a lo largo de la progresión y transformación tumoral<sup>133-135,143,185</sup>. Otros estudios han demostrado que las mutaciones del gen *CREBBP* en DLBCL pueden dirigir la patogénesis de la enfermedad a través de la desregulación de la función de *BCL6* o *TP53*<sup>132</sup>. Así, en células B normales, la acetilación de *BCL6* mediada por *CREBBP* conduce a la inactivación de su función de transrepresión, mientras que la misma modificación postraduccional representa un requisito esencial para la activación

de las funciones supresoras de tumores de *p53*. Este equilibrio se interrumpe en el DLBCL, debido a la presencia de mutaciones y/o deleciones en *CREBBP*. Así, las mutaciones en *CREBBP* parecen contribuir a la linfomagénesis favoreciendo la reducción de la actividad de un supresor tumoral como es *p53* y produciendo la activación constitutiva del oncogén *BCL6*<sup>132</sup>.

En conjunto, estas observaciones indican que las mutaciones de *CREBBP* desempeñan un papel fundamental en el desarrollo del FL y DLBCL, y los efectos fisiológicos a los que dan lugar pueden ser dirigidos por un mecanismo de acetilación desregulada de las histonas y/o de proteínas no histonas, que podría favorecer la reprogramación epigenética y, así, escapar la enfermedad a la inmunidad celular. Por ello, uno de los trabajos que conforman esta tesis doctoral está enfocado a determinar qué función cumplen las mutaciones de pérdida de función del gen *CREBBP* en el desarrollo del linfoma folicular y del linfoma difuso de células B grandes, dado que en la actualidad no está claro **i)** la célula de origen de dicho linfoma, y **ii)** si las consecuencias funcionales de dichas mutaciones son las mismas en estas dos enfermedades. Además, el papel de estos eventos en la linfomagénesis de las células B sólo se ha estudiado en líneas celulares sin utilizar modelos de ratón genéticamente modificados.

## 5. *LMO2* y su función en el desarrollo de la leucemia linfoblástica aguda de células T.

### 5.1 *Leucemia linfoblástica aguda de células T: definición, epidemiología, clínica, clasificación.*

La leucemia linfoblástica aguda de células T (T-ALL) es un tumor linfoide inmaduro caracterizado por la infiltración difusa de la médula ósea por células hematopoyéticas malignas que expresan marcadores de células T. Las T-ALL representan entre el 10-15% de los casos pediátricos<sup>186,187</sup> y entre el 20-25% de los casos adultos<sup>188,189</sup> de leucemia linfoblástica aguda (ALL) en Europa, Estados Unidos y Japón, siendo dos veces más prevalente en hombres que en mujeres.

Clínicamente, los pacientes con T-ALL suelen presentar recuentos de glóbulos blancos elevados en su sangre e insuficiencia hematopoyética, con neutropenia, anemia y trombocitopenia. Además, con frecuencia, presentan masas mediastínicas tónicas e infiltración meníngea del sistema nervioso central en el momento del

diagnóstico. Los tumores inmaduros de células T que se presentan como una masa tímica pero solo infiltran la médula ósea, se diagnostican como linfoma linfoblástico de células T.

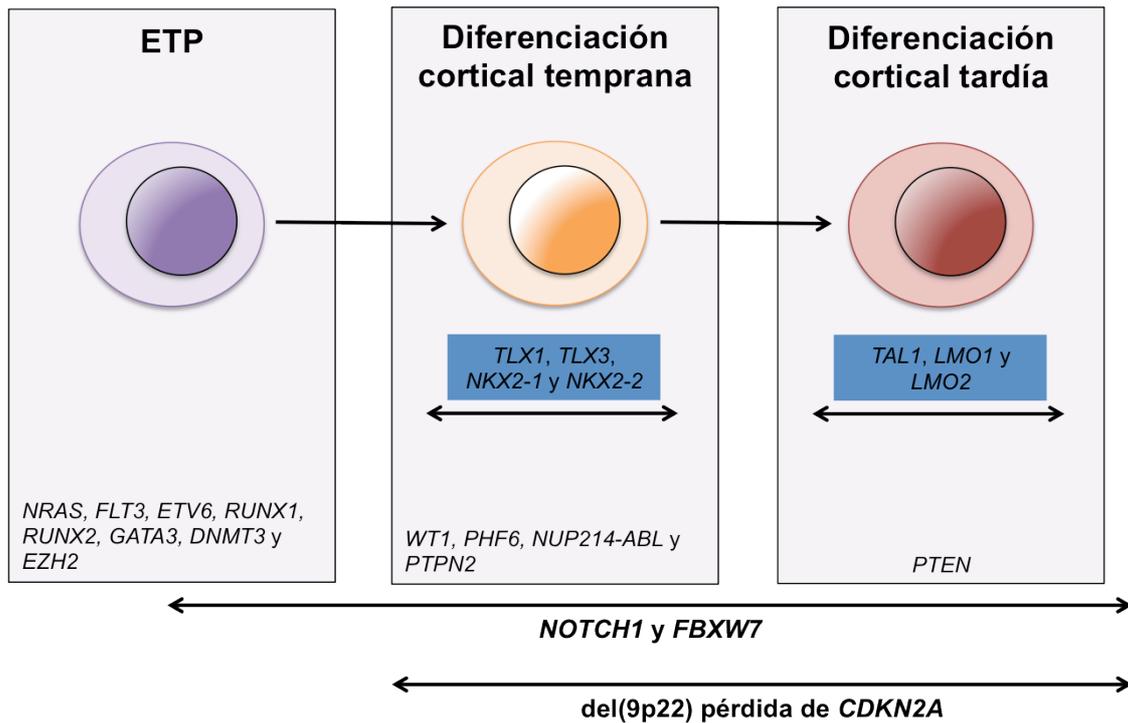
En la era de la quimioterapia combinada, los resultados para los pacientes pediátricos de ALL mejoraron notablemente, pero el 20% de estos pacientes todavía mueren de la enfermedad y este número aumenta al 50% en los adultos<sup>189</sup>. Cuando se usó por primera vez la quimioterapia combinada en la clínica, se reconoció a los pacientes con T-ALL como un subgrupo de leucemia de alto riesgo con unas tasas de curación del 10% que nada tenían que ver con la tasa del 40% que se da en la leucemia linfoblástica aguda de células B precursoras<sup>190</sup>. Esta observación clínica condujo a la introducción de protocolos intensificados de quimioterapia en el tratamiento de las T-ALL y como consecuencia se produjo una mejora gradual en los resultados. Las tasas actuales de curación en ensayos multicéntricos son del 80% en niños<sup>191</sup> y del 60% en adultos<sup>191</sup>. A pesar de este progreso, el pronóstico de los pacientes con T-ALL resistente, que no logran obtener una remisión hematológica completa, o aquellos cuya enfermedad recidiva después de una respuesta inicial transitoria, sigue siendo bastante desesperanzador<sup>192</sup>. Frente a estos desafíos clínicos, se han dedicado muchos esfuerzos para descifrar los eventos moleculares subyacentes a la transformación de la T-ALL, con el objetivo de identificar dianas terapéuticas más específicas y desarrollar fármacos y combinaciones de fármacos más eficaces y menos tóxicos<sup>193</sup>.

Todos los esfuerzos que se han llevado a cabo para determinar los mecanismos moleculares que subyacen a las T-ALL han permitido clasificar a las T-ALL, no sólo por sus características clínicas, sino también por sus firmas de expresión genética<sup>194</sup> y su inmunofenotipo. De esta manera podemos clasificar a las T-ALL en tres subgrupos:

- 1. Leucemias de progenitores tempranos de linaje T (ETP):** se trata de un tipo de leucemias que tienen un bloqueo en las etapas más tempranas de la diferenciación de células T (células CD4<sup>-</sup> CD8<sup>-</sup>) y muestran un programa transcripcional asociado no sólo con los progenitores tempranos de linaje T, sino también asociado con células madre hematopoyéticas y progenitores mieloides<sup>195,196</sup>. Genéticamente, las ETP T-ALL tienen una menor prevalencia de mutaciones en *NOTCH1*, rara vez presentan deleciones en *CDKN2A* y están característicamente asociadas con mutaciones en genes que codifican factores de la señalización (como por ejemplo *NRAS* y *FLT3*) y reguladores

epigenéticos (como por ejemplo *IDH1*, *IDH2*, *DNMT3* y *EZH2*)<sup>197,198</sup>. Este tipo de leucemias también se asocian con mutaciones que alteran la actividad de importantes factores de transcripción encargados de dirigir el desarrollo hematopoyético y de células T (por ejemplo, *RUNX1*, *GATA3* y *ETV6*)<sup>197,198</sup> (**Figura 5**). Aunque este tipo de leucemias representan aproximadamente el 10% de todos los casos de T-ALL pediátricos<sup>195</sup>, su incidencia aumenta considerablemente con la edad, representando aproximadamente el 40-50% de las T-ALL en adultos<sup>196</sup>. Se trata de un grupo de leucemias con un alto riesgo de recaídas y baja supervivencia tanto en niños como en adultos.

- 2. T-ALL de timocitos corticales en estadios tempranos de maduración:** se trata de un tipo de T-ALLs en las que se produce un bloqueo en las primeras etapas de la maduración cortical del linaje T. Presentan un inmunofenotipo  $CD1a^+CD4^+CD8^+$ . Este tipo de leucemias muestran un pronóstico particularmente favorable<sup>199</sup>. Desde un punto de vista molecular, estas leucemias se asocian con la activación de genes con dominio *homeobox*, entre los que encontramos a *TLX1*, *TLX3*, *NKX2-1* y *NKX2-2*; además, tienen la más alta prevalencia de mutaciones en *NOTCH1* y de deleciones del locus de *CDKN2A*<sup>200,201</sup>. También suelen presentar alteraciones genéticas características que rara vez aparecen en los otros subtipos de T-ALL, entre las que encontramos reordenamientos de *NUP214-ABL*<sup>202</sup>, y mutaciones en *PHF6*<sup>203</sup>, *WT1*<sup>204</sup> y *PTPN2*<sup>205</sup> (**Figura 5**).
- 3. T-ALL de timocitos corticales en estadios tardíos de maduración:** se trata del tercer subgrupo de T-ALL, en las que se produce un bloqueo en la diferenciación de los timocitos corticales en etapas más maduras. En este caso, el inmunofenotipo característico es  $CD3^+CD4^+CD8^+$ . A nivel molecular, estas leucemias suelen mostrar activación de factores de transcripción como *TAL1*<sup>206</sup>, *LMO1*<sup>207</sup> o *LMO2*<sup>208</sup>. Además, suelen presentar alteraciones secundarias que con frecuencia afectan a *NOTCH1*<sup>209</sup>, *FBXW7*<sup>210</sup> y *CKN2A*<sup>211</sup>, pero también una mayor prevalencia de mutaciones y deleciones de *PTEN*<sup>212</sup> (**Figura 5**).



**Figura 5: Los programas oncogénicos y la huella de expresión génica definen los diferentes subgrupos moleculares de T-ALL.** Adaptado de L. Belver and A. Ferrando. *The genetics and mechanism of T-ALL. Nature Reviews Cancer.* 2016<sup>192</sup>.

Como hemos comentado anteriormente, la etiología de las neoplasias malignas hematológicas sigue siendo difícil de entender y es objeto de intensas investigaciones. El aislamiento de las células tumorales primitivas / iniciadoras ha sido crítico para mejorar nuestra comprensión del origen y la patobiología de la leucemia. Además, la heterogeneidad clínica, molecular y biológica de la leucemia apunta a una diversidad multiclonal, hasta ahora subestimada, que pone de relieve los orígenes complejos y multifactoriales de las leucemias<sup>213</sup>. De acuerdo con la teoría jerárquica del desarrollo del cáncer que implica a las células madre del cáncer (CSC) como fuente de procesos malignos<sup>214,215</sup>, y la hipótesis de múltiples eventos, la leucemia debe ser el resultado de una célula clonal normal afectada por una primera señal oncogénica que la transforma en un precursor premaligno "reprogramado" que está predispuesto a sufrir transformaciones adicionales debido a translocaciones y/o mutaciones de genes<sup>213,216</sup>. En consonancia con esto, la participación de células primitivas en la etiología tanto de las B-ALL<sup>36,217,218</sup> como de las T-ALL<sup>37</sup> ha sido sugerida. Sin embargo, la contribución de una única célula de origen en la ALL ha sido fuertemente debatida, ya que no siempre es aparente<sup>219</sup>. Además, la identidad exacta de la célula de origen en las T-ALL es menos clara. Tradicionalmente, se ha pensado

que el origen celular de la T-ALL son progenitores linfoides comprometidos al linaje T, arrestados en el mismo estadio celular que la masa de células tumorales, de manera que, el fenotipo tumoral refleja la célula de origen de la leucemia. Sin embargo, en la actualidad son muchos los que opinan que dicha célula de origen de las T-ALL no tendría por qué ser necesariamente una célula comprometida al linaje T.

Dada la pobre evidencia clínica de que las T-ALL presenten una organización jerárquica, los modelos murinos han proporcionado poderosos conocimientos moleculares sobre la leucemogénesis de las T-ALL y ponen en evidencia la necesidad de determinar cual es su verdadera célula de origen y si un mecanismo de reprogramación tumoral estaría dirigiendo el inicio de las mismas. Con el fin de responder a estas preguntas centramos el segundo objetivo de esta tesis doctoral en el estudio del papel del oncogén *LMO2* en la génesis de las T-ALL.

### **5.2 La proteína *LMO2* humana: descubrimiento, estructura, función y patogénesis.**

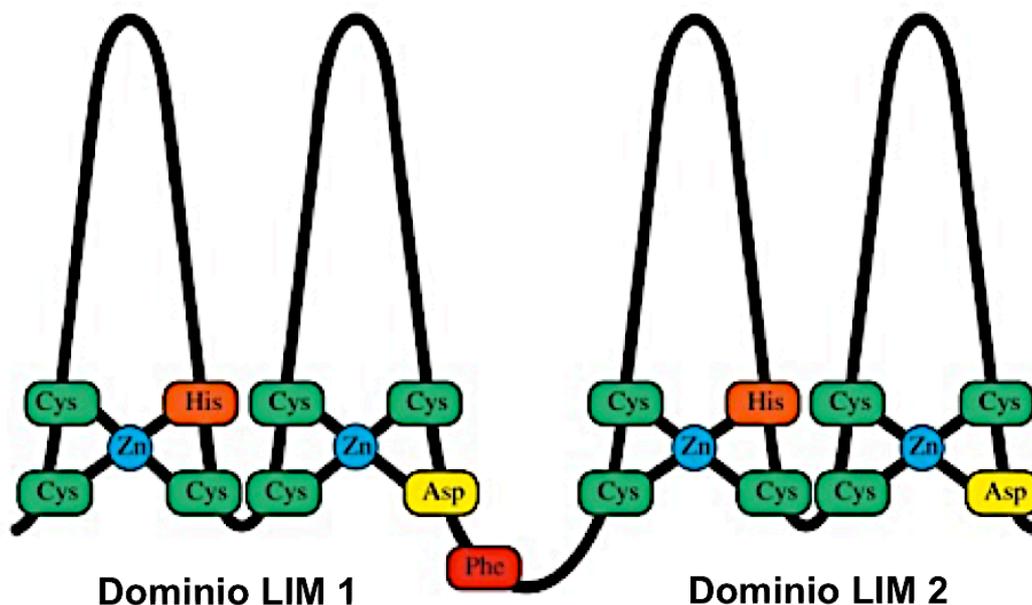
Los avances en biología molecular de los años 80 condujeron al hallazgo de las translocaciones cromosómicas recurrentes asociadas al cáncer, que dan lugar a un control perturbado de la expresión del oncogén como consecuencia de la unión del mismo con los genes de las Ig o con el receptor de células T (TCR), o bien dan lugar a la creación de una nueva proteína de fusión con funciones quiméricas<sup>220,221</sup>. Estudios de reordenamientos del gen del receptor de células T (*TCR*), sugirieron que el cáncer de células T asociado a translocaciones cromosómicas tenía activaciones oncogénicas<sup>222</sup>. Esta hipótesis finalmente se demostró y, *LMO2* se descubrió hace ya 27 años, publicándose un año después de su descubrimiento como este gen estaba asociado con una translocación cromosómica recurrente del loci del *TCR* en un subgrupo de pacientes con leucemia linfoblástica aguda de células T (T-ALL)<sup>223,224</sup>. Desde entonces, los estudios sobre esta proteína han demostrado que *LMO2* está implicado en multitud de efectos celulares, que van desde un papel proto-oncogénico en células T, a un papel esencial en la hematopoyesis y la remodelación vascular, así como un importante papel en la biología de las células madre<sup>225</sup>.

El primer gen descubierto de la familia a la que pertenece *LMO2* fue *LMO1* (anteriormente conocido como *RBTN1*, *Rombotina-1* o *Ttg-1*). *LMO1* fue uno de los primeros proto-oncogenes que se aislaron asociados a las translocaciones recurrentes en las T-ALL<sup>226-228</sup>. Además, estos trabajos sugirieron que otros proto-oncogenes similares podían existir dentro del genoma, y así, *LMO2* (anteriormente conocido como

*RBTN2*, *Rombotina-2* o *Ttg-2*) fue descubierto posteriormente por dos laboratorios independientes, el primero, a través de experimentos de hibridación con el objetivo de encontrar genes homólogos de *LMO1*<sup>223</sup> y, el segundo, por clonación directa de los puntos de ruptura de la translocación t(11,14)(p13; q11)<sup>224</sup>. Así, se confirmó que *LMO2*, al igual que *LMO1*, se encuentra en el brazo corto del cromosoma 11 humano, pero en la banda 11p13 en lugar de en la banda 11p15 (donde se localiza *LMO1*).

La familia de proteínas LMO es ahora conocida por contener 4 genes, *LMO1*, 2, 3 y 4<sup>229-232</sup>. El nombre de esta familia de proteínas procede de su estructura terciaria, que se compone de dos regiones dispuestas en tándem llamadas dominios LIM (dominio LIM1 y LIM2). Se trata de dominios ricos en cisteínas, que se caracterizan por una secuencia de aminoácidos consenso  $CX_2-CX_{17-19}-HCX_2-CX_2-CX_{16-20}-CX_2CX/C/D$ <sup>233</sup> (**Figura 6**). *LMO1*, *LMO2* y *LMO3* están involucrados en las translocaciones cromosómicas que aparecen en las T-ALL, mientras que *LMO4* es el miembro más divergente de la familia y actualmente no se asocia con translocaciones conocidas en leucemia. Los fenotipos de los cuatro genes en ratones *knock-out* muestran su papel crucial en las decisiones de destino celular, y hay sinergia fenotípica cuando la pérdida homocigótica de *Lmo1* y *Lmo3* se produce en ratones<sup>234,235</sup>. Estas propiedades fenotípicas, llevaron, en parte, a la hipótesis de “los genes maestro”<sup>220</sup>, basada en la importancia para el desarrollo, de algunos activadores transcripcionales<sup>220,236</sup>. Así, *LMO2* sería un ejemplo arquetipo de un regulador transcripcional maestro, pero también de un oncogén asociado a translocación cromosómica.

*LMO2*, por tanto, es una proteína de dominio LIM, en la que los dos dedos LIM se pliegan a través de la unión de un átomo de zinc por las cadenas laterales de tres residuos de cisteína y un residuo de histidina (dedo 1), o tres residuos de cisteína y un residuo de ácido aspártico (dedo 2)(**Figura 6**). Los dedos LIM son distintos de los dedos de zinc de los factores de transcripción, como los de la familia GATA, ya que no existen evidencias de dominios LIM que se unan directamente a ADN. Más bien, estos dedos LIM permiten la interacción de *LMO2* con otras proteínas asociadas<sup>237-240</sup>. Cada dedo está separado por dos residuos de aminoácidos, cuya separación y secuencia se ha demostrado que es vital para las subsiguientes interacciones de esta proteínas con otras<sup>241</sup>.



**Figura 6: Diagrama esquemático de las proteínas LMO que muestra la disposición en tándem de los dos dominios LIM.** Cada dominio comprende dos estructuras de tipo dedo de zinc, que coordinan un átomo de zinc entre cuatro residuos. Los dedos de cada dominio están unidos por dos residuos de aminoácidos que se conservan entre especies. Adaptado de I. Sánchez-García and T.H. Rabbits. *The LIM domain: a new structural motif found in zinc-finger-like proteins. Trends in Genetics.* 1994<sup>228</sup>.

Así, LMO2 no se une directamente al ADN, sino que actúa formando parte de un puente para que las proteínas de unión al ADN creen un complejo bipartito de unión al ADN (**Figura 7**). La proteína LMO2 recién sintetizada no tiene una conformación estable, lo que requiere de la interacción con otras proteínas con el objetivo de estabilizar su estructura y evitar la degradación proteosómica. Una de las proteínas que interactúan con LMO2 para estabilizarlo es TAL1<sup>242</sup>. LMO2 se mostró, por primera vez, formando parte de un complejo formado por TAL-1, E47 y GATA-1 en células eritroides<sup>238</sup>. Posteriormente, se descubrió que LMO2 también interactuaba con otras proteínas como LYL1 y TAL2<sup>243</sup>. Además, LMO2 también puede unirse a las proteínas GATA-1, GATA-2 o GATA-3<sup>240,244,245</sup>. El complejo LMO2 contiene además una proteína adicional llamada LDB1 (dominio LIM de unión a proteína 1)<sup>240,246</sup>. Así, el complejo LMO2 está formado por un complejo bipartito de unión al ADN que reconoce un sitio E-box y GATA, separados por una vuelta de doble hélice de ADN (**Figura 7**). De esta manera el complejo de LMO2 actúa activando la transcripción de diferentes genes diana.

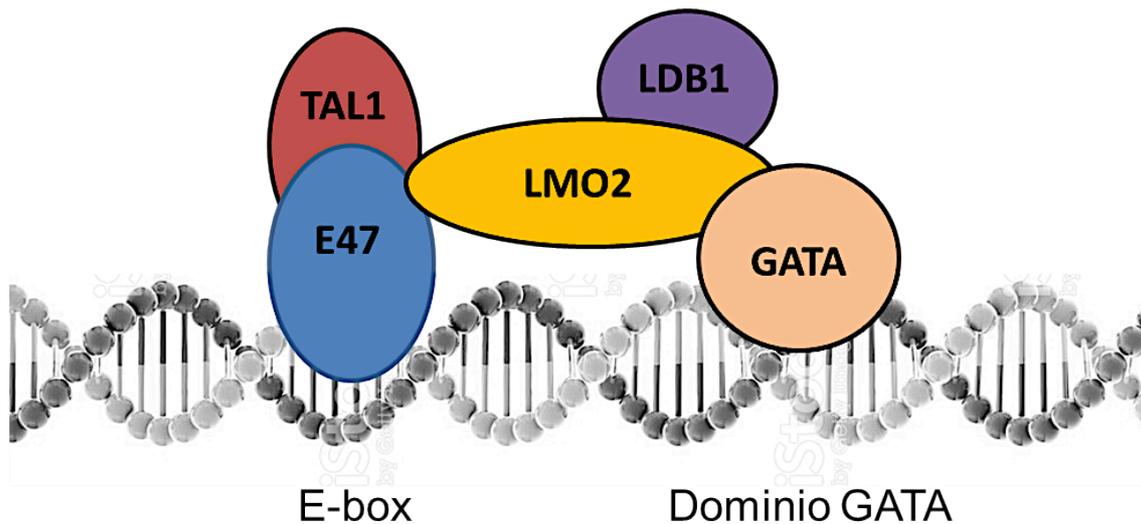


Figura 7: Complejo pentamérico de LMO2, comúnmente encontrado en células de linaje eritroide, unido a un dominio E-box-GATA bipartito.

Tras la identificación de *LMO2* como un proto-oncogén en algunas leucemias linfoblásticas agudas de células T, la función natural de la proteína LMO2 se estudió en células de ratón y humanas con el fin de determinar la bioquímica normal de la misma. Un primer hallazgo determinó que la expresión de *LMO2* en células hematopoyéticas tenía una función crucial en la especificación del linaje hematopoyético. La inserción del gen de la  $\beta$ -galactosidasa como reportero de la expresión de *Lmo2* en ratones, mostró que la expresión de *Lmo2* se daba principalmente en el sistema hematopoyético y en el desarrollo del cerebro<sup>247</sup>. En la jerarquía hematopoyética, *LMO2* tiene un papel crucial en el desarrollo de los glóbulos rojos en el saco vitelino embrionario (eritropoyesis primitiva), de ahí que los ratones nulos para *Lmo2* mueran en el útero en la etapa embrionaria E9-E10<sup>234</sup>. Curiosamente, se demostró que las consecuencias de la pérdida homocigota de *Lmo2* eran similares a las encontradas en los embriones nulos para *Tal1*<sup>248,249</sup>, *Gata-1*<sup>250,251</sup> y *Gata-2*<sup>252</sup>, presagiando que, como indicamos anteriormente, la proteína LMO2 formaba parte de un complejo pentamérico que incluía a las proteínas TAL1 y GATA<sup>240,243,244</sup>.

Dado que los embriones nulos para *Lmo2* no sobreviven a término, se requirió una estrategia diferente para evaluar el papel de esta proteína en la hematopoyesis definitiva. La hematopoyesis definitiva se inicia en el embrión a la mitad de la gestación, en la región aorta-gónada-mesonefros<sup>253,254</sup>, cambia al hígado fetal y se

traslada finalmente a la médula ósea poco antes del nacimiento, donde persiste durante toda la vida adulta. Por consiguiente, para el estudio del papel de *LMO2* en el desarrollo hematopoyético, se usaron células ES con mutación nula homocigota para *Lmo2* con el objetivo de crear ratones quiméricos que presentaran una contribución tanto de las células del blastocisto receptor como de las células donantes ES nulas para *Lmo2*. Estos estudios mostraron que las células ES nulas para *Lmo2* no contribuían a la hematopoyesis de ratones adultos<sup>234,255</sup>, lo que sugiere que la función de *LMO2* es crucial a nivel de las células madre hematopoyéticas multipotentes (HSCs). Además, también se demostró que la sobreexpresión de *LMO2* (y *LMO1*) en líneas celulares de eritroblastos, daba lugar a células precursoras inmaduras de linaje eritroide, lo que sugería que *LMO2* podría actuar como un regulador negativo de la diferenciación<sup>256</sup>.

*LMO2* también se expresa en células madre hematopoyéticas, encargándose de su mantenimiento, donde parece formar complejos proteicos con *GATA2* y *LYL1*<sup>257</sup>. Estudios recientes involucran a un grupo de siete reguladores transcripcionales, entre los que se encuentra *LMO2*, en el control del mantenimiento y diferenciación de las células madre hematopoyéticas: *SCL*, *LYL1*, *LMO2*, *GATA2*, *RUNX1*, *ERG* y *FLI-1*<sup>258</sup>. Además, se ha comprobado que *Lmo2* está involucrado en el proceso de reprogramación hacia pluripotencia de células hematopoyéticas diferenciadas<sup>259</sup>.

Por lo tanto, *LMO2* tiene un papel crucial a nivel de las células madre hematopoyéticas y además puede conferir propiedades de célula madre en células diferenciadas.

También se ha demostrado la participación de *Lmo2* en otros procesos biológicos como la remodelación angiogénica de las redes vasculares<sup>247</sup>, la vascularización *de novo* en tumores<sup>260</sup>, y la formación de células endoteliales<sup>261</sup>. Por otro lado, *Lmo2* se asocia al desarrollo de tejidos del sistema nervioso central (SNC)<sup>229</sup>, expresándose en el cerebro del ratón adulto, junto con *Lmo1* y *Lmo3*<sup>262</sup>, en particular después de las crisis epilépticas, posiblemente como un mecanismo de reparación.

A pesar de que *LMO2* está asociado con leucemias de células T, no es necesario para el desarrollo linfóide T, expresándose sólo en etapas tempranas de la diferenciación T (en el estadio doble negativo para CD4 y CD8) y no expresándose en estadios diferenciados del linaje T<sup>263,264</sup>.

En cuanto a la patogénesis de *LMO2*, tal y como se comentó al principio de esta sección, *LMO2* fue descubierto asociado a la translocación cromosómica t(11,14)(p13;

q11)<sup>224</sup> que aparece en las leucemias linfoblásticas agudas de células T humanas y que tiene como resultado una expresión aberrante del gen. De ahí que, con el objetivo de estudiar el papel de este gen en la génesis de las T-ALL<sup>208,265</sup>, se desarrollaran modelos de ratón transgénicos en los que se indujo la expresión de *Lmo2* bajo el control de diferentes promotores específicos de células diferenciadas o progenitores comprometidos de linaje T: *CD2*<sup>266-268</sup> o *Lck*<sup>269</sup>. El resultado de todos estos modelos fue el desarrollo de neoplasias de células T clonales y de latencia prolongada. Según estos estudios, *Lmo2* ejerce su efecto oncogénico en las células T inmaduras en el timo, para posteriormente inducir neoplasias de células T con evidentes reordenamientos clonales del gen *TCR*. Por lo tanto, y atendiendo a estos estudios, *Lmo2* es necesario, pero no suficiente para inducir el desarrollo de leucemia, siendo necesaria la aparición de mutaciones secundarias. Un tipo de mutación que se encuentra con alta frecuencia en las neoplasias T de estos ratones es la mutación de *Notch1*<sup>269</sup>, mutación observada en más del 50% de las T-ALL humanas<sup>270</sup>. Por lo tanto, todos estos estudios van orientados a inducir la expresión de *Lmo2* en células T del timo, con el objetivo de inducir T-ALL similares a las humanas, dado que hasta el momento, existe la idea de que este tipo de leucemias surgen en el timo. Sin embargo, la aparición de T-ALL en niños con síndrome de inmunodeficiencia combinada grave ligada al cromosoma X (SCID-X1), tratados con terapia génica, puso en entredicho esta idea de que las T-ALL tienen como célula de origen una célula T del timo.

Por otro lado, aunque las anomalías cromosómicas que implican a *LMO2* parecen estar restringidas a tumores de células T, estudios de expresión génica han implicado a *LMO2* en otra serie de tumores. La expresión de *LMO2* se ha observado en las células B del centro germinal normal y en el linfoma difuso de células B grandes derivadas del centro germinal<sup>271,272</sup>, donde la presencia de *LMO2* es un factor de buen pronóstico<sup>273</sup>. Además, la expresión de *LMO2* también se ha detectado en algunas leucemias agudas de células B (B-ALL)<sup>274-276</sup> y, en este caso, presenta correlaciones pronósticas diferentes, existiendo una buena asociación pronóstica en algunos casos<sup>275</sup>, pero en otros, como las B-ALL asociadas al gen de fusión *E2A-HLF*, la expresión de *LMO2* es una característica del mal pronóstico<sup>274,276</sup>. En el caso de los tumores sólidos, la expresión de *LMO2* también es un factor de buen pronóstico en los pacientes con cáncer de páncreas<sup>277</sup>.

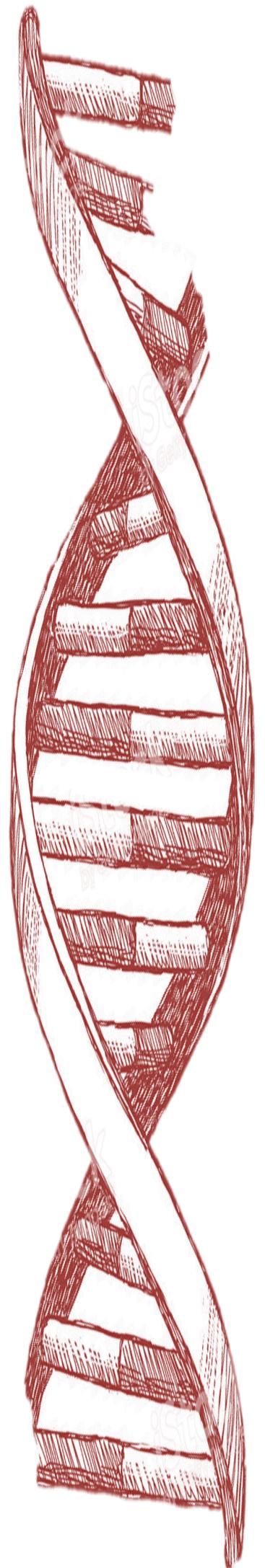
### 5.3 LMO2 y terapia génica.

A finales de 1990 y principios de 2000, se iniciaron ensayos de terapia génica que iban dirigidos a lactantes con inmunodeficiencia combinada grave ligada al cromosoma X (SCID-X1), una inmunodeficiencia primaria severa causada por mutaciones inactivantes del receptor y de la interleucina-2 (*IL2RG*)<sup>278,279</sup>. De los 20 recién nacidos tratados, la mayoría mostró mejoras en el número y función de células T y células inmunes, pero cinco desarrollaron leucemia entre 2-6 años después del tratamiento<sup>280-282</sup>. En cuatro casos, el vector retroviral que contenía el gen del *IL2RG* corregido se había insertado cerca del dominio del gen *LMO2*, dando lugar a la sobreexpresión de *LMO2* (en uno de estos casos se produjo una inserción adicional cerca del proto-oncogén *BMI1*). Este desafortunado efecto secundario de los ensayos de terapia génica, sin embargo, proporcionó un experimento no intencional, pero imparcial, sobre la mutagénesis de la inserción retroviral humana y planteó varias preguntas sobre por qué se observó la inserción recurrente en *LMO2*, pero no en otros miembros de la familia LMO o en otros oncogenes asociados a la T-ALL, como *TAL1*. La alta tasa de inserción en un solo gen (*LMO2*) fue sorprendente, y una investigación posterior indicó que varios factores podían contribuir a este fenómeno. Los retrovirus preferentemente se integran cerca del extremo 5' de los genes transcritos activamente<sup>283</sup>, por lo que la expresión de *LMO2* en la hematopoyesis hace que sea un objetivo para la mutagénesis por inserción retroviral, pero no explica por qué otros genes expresados tempranamente no fueron dirigidos de manera similar. Como una posible explicación, se propuso la existencia de un sitio "frágil" alrededor de *LMO2* que haría que esta región fuera propensa tanto a los puntos de ruptura cromosómicos como a la integración del vector retroviral<sup>284</sup>. También podría existir alguna cooperación entre *IL2RG* y *LMO2* en la inducción de la leucemia, aunque esta cooperación requeriría mutaciones adicionales antes de dar lugar a la enfermedad<sup>285,286</sup>. Sin embargo, se ha debatido sobre si *IL2RG* contribuye activamente a la leucemogénesis<sup>287</sup>. Estos problemas no están completamente resueltos, y sugieren que *LMO2* es capaz de reprogramar a una célula madre hematopoyética hacia un linaje T tumoral dentro de las T-ALL humanas. Sin embargo, esta capacidad de reprogramación tumoral de una HSC por parte del gen *LMO2*, en las T-ALL, aún no ha sido demostrada experimentalmente.





*Hipótesis de trabajo*





### HIPÓTESIS DE TRABAJO

El cáncer es una enfermedad clonal, que tiene su origen en una célula específica, llamada célula de origen del cáncer. Hasta el momento, la mayoría de estudios publicados han centrado sus esfuerzos en determinar el papel de los oncogenes en el desarrollo del cáncer, pero muy pocos se han centrado en estudiar cómo se establece la identidad tumoral. En general, se piensa que el cáncer surge en una célula con identidad prefijada, que es la responsable del fenotipo tumoral. Además, este modelo actual del cáncer sostiene que, en los cánceres humanos, la función de los oncogenes es homogénea en todas las células cancerosas, con independencia de la heterogeneidad celular existente dentro del tumor. Desafortunadamente, las terapias basadas en este modelo de cáncer no logran erradicar los tumores humanos, sugiriendo que la función de los oncogenes en la masa tumoral no es homogénea.

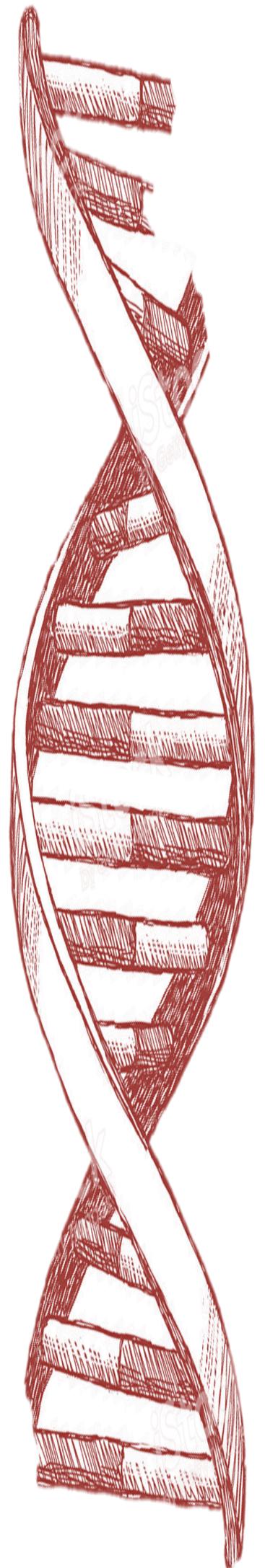
Por el contrario, nuestra hipótesis sostiene que el cáncer surge como consecuencia de un mecanismo de reprogramación tumoral de la célula de origen<sup>7</sup>. Este mecanismo modificaría el estado epigenético de la célula iniciadora del cáncer, estableciendo en ella un programa de diferenciación celular patológico. Así, como consecuencia de la escasez de estudios sobre la célula iniciadora del cáncer, no se han desarrollado modelos animales que reproduzcan fielmente la enfermedad humana.

Por ello, en este trabajo de tesis doctoral nos planteamos determinar si el linfoma de células B asociado a alteraciones de *CREBBP* y la leucemia linfoblástica aguda de células T vinculada a *LMO2* son consecuencia de un mecanismo de reprogramación tumoral de la célula de origen y cuál es la célula donde se originan.





*Objetivos*





### OBJETIVOS

#### **PRIMERO - Papel de *CREBBP* en la génesis del linfoma de células B.**

- Definir el papel de la célula de origen en la linfomagénesis de células B mediada por *CREBBP*.
- Determinar qué función cumplen las mutaciones de pérdida de función del gen *CREBBP* en el desarrollo del linfoma folicular y del linfoma difuso de células B grandes, estudiando si las consecuencias funcionales de dichas mutaciones son las mismas en estas dos enfermedades.

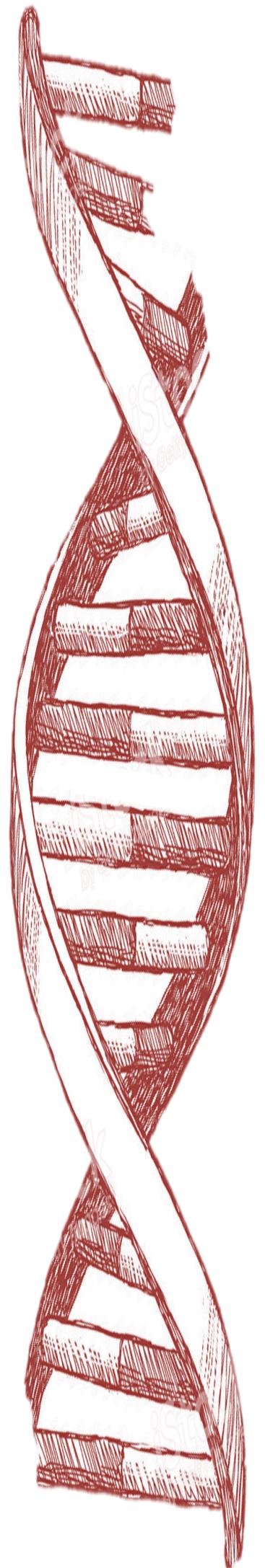
#### **SEGUNDO - Papel de *LMO2* en la génesis de la leucemia linfoblástica aguda de células T.**

- Estudiar si la expresión limitada de *Lmo2* al compartimento de células madre/progenitoras hematopoyéticas es suficiente para dar lugar a una leucemia linfoblástica aguda de células T como consecuencia de un mecanismo de reprogramación tumoral.
- Localizar el nicho celular donde están teniendo lugar los eventos genéticos secundarios asociados a la leucemia linfoblástica aguda de células T.





*Resultados*





## RESULTADOS

### 1. La pérdida de *Crebbp* a nivel de células pro-B coopera con la sobreexpresión de *Bcl2* para promover la linfomagénesis.

#### 1.1 La pérdida de *Crebbp* a nivel de célula pro-B produce una alteración del desarrollo normal de las células B.

Para investigar el papel de *Crebbp* en el desarrollo de células B y la célula de origen de la linfomagénesis a la que se asocia, generamos diferentes líneas de ratones transgénicos (**Tabla II**). *CREBBP* se caracteriza por presentar una inactivación monoalélica y bialélica en tumores humanos. Por ello, obtuvimos cepas con uno o ambos alelos del gen *Crebbp* flanqueados por sitios *LoxP* (alelos a los que llamaremos condicionales o *F*) con el objetivo de mimetizar al máximo la enfermedad humana. Dichos ratones fueron cedidos por el laboratorio de *Paul Brindle*<sup>288</sup>. Además, se piensa que las mutaciones de *CREBBP* surgen como un evento secundario temprano en el FL<sup>143,185</sup>, lo que nos hizo pensar que las mutaciones de *CREBBP* podrían estar ocurriendo en una etapa temprana de la diferenciación B en estos linfomas. Por ello, cruzamos los animales que presentaban uno o dos alelos de *Crebbp* condicionales con la cepa de ratones *mb1-cre*<sup>289</sup>, que expresa la enzima Cre-recombinasa de manera específica a nivel de células pro-B, con el fin de eliminar el(los) alelo(s) condicional(es) de *Crebbp* de manera específica en estadios tempranos de la diferenciación B. Así, generamos los ratones *Cbp*<sup>WT/ $\Delta$</sup>  (con un solo alelo de *Crebbp* deletado) y *Cbp* <sup>$\Delta/\Delta$</sup>  (con ambos alelos del gen deletados).

Varios estudios han descrito la presencia de translocaciones del gen *BCL2* que cooperan con las mutaciones de *CREBBP* tanto en pacientes con linfoma folicular como en aquellos con DLBCL de células B del centro germinal<sup>131, 143</sup>. Además, se piensa que la translocación de *BCL2* en el linfoma humano se adquiere a nivel de una célula pre o pro-B<sup>290</sup>. Por ello, con el objetivo de diseñar un modelo que reprodujera lo más fielmente la enfermedad humana, decidimos cruzar la cepa *E $\mu$ Bcl2*, que sobreexpresa el oncogén *Bcl2* en las células B del ratón bajo el control del *enhancer* de la cadena pesada de las inmunoglobulinas (IgH)<sup>120</sup>, con los ratones *Cbp*<sup>WT/ $\Delta$</sup>  y *Cbp* <sup>$\Delta/\Delta$</sup> , a fin de producir ratones que inactivaran uno o dos alelos de *Crebbp* y sobreexpresaran *Bcl2* en células B. Como cepas control del estudio utilizamos ratones que presentaban uno o dos alelos de *Crebbp* condicionales pero sin deletar (cepas *Cbp*<sup>WT/*F*</sup> y *Cbp*<sup>*F*/*F*</sup>), ratones que sólo sobreexpresaban *Bcl2* a nivel de células B (*E $\mu$ Bcl2*) y ratones que sobreexpresaban *Bcl2* a nivel de células B y además presentaban ambos alelos de *Crebbp* condicionales (cepa *Cbp*<sup>*F*/*F*</sup> x *E $\mu$ Bcl2*).

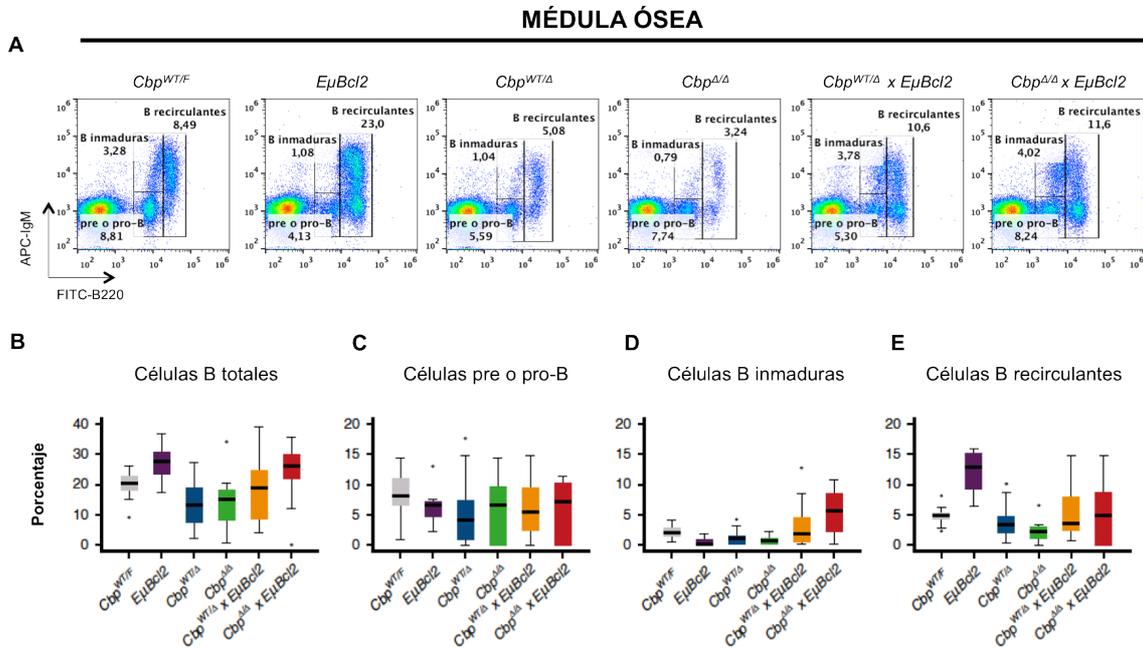
**Tabla II: Cepas de ratón transgénicas utilizadas para investigar el papel de la inactivación de *Crebbp*.**

Cepa de ratón	<i>Crebbp</i> condicional	<i>mb1-cre</i>	<i>EuBcl2</i>	Fenotipo en célula B
<i>Cbp<sup>WT/F</sup></i>	-/+	-	-	Sin efecto
<i>Cbp<sup>F/F</sup></i>	+/+	-	-	Sin efecto
<i>EuBcl2</i>	-/-	-	+	Sobreexpresión de <i>Bcl2</i>
<i>Cbp<sup>F/F</sup> x EuBcl2</i>	+/+	-	+	Sobreexpresión de <i>Bcl2</i>
<i>Cbp<sup>WT/Δ</sup></i>	-/+	+	-	Un alelo de <i>Crebbp</i> deletado
<i>Cbp<sup>Δ/Δ</sup></i>	+/+	+	-	Dos alelos de <i>Crebbp</i> deletados
<i>Cbp<sup>WT/Δ</sup> x EuBcl2</i>	-/+	+	+	Un alelo de <i>Crebbp</i> deletado y sobreexpresión de <i>Bcl2</i>
<i>Cbp<sup>Δ/Δ</sup> x EuBcl2</i>	+/+	+	+	Dos alelos de <i>Crebbp</i> deletados y sobreexpresión de <i>Bcl2</i>

Así, estudiamos el desarrollo de células B de ratones jóvenes (de 11-20 semanas de edad) y de ratones adultos de más de 5 meses de edad, usando citometría de flujo y comparando todas las cepas entre sí (**Anexo 1 y Anexo 2**, respectivamente). Aquellas cepas con *Crebbp* inactivado mostraron un déficit en los porcentajes de diferentes subpoblaciones de células B de la médula ósea y el bazo (**Figura 8 y Figura 9**), apoyando la idea de que *Crebbp* es un importante regulador de la diferenciación de células B<sup>288</sup>. Los cambios en los porcentajes de células B incluyeron un número significativamente reducido de células B B220<sup>+</sup> tanto en la médula ósea como en el bazo de los ratones *Cbp<sup>WT/Δ</sup>* y *Cbp<sup>Δ/Δ</sup>* en comparación con los ratones *Cbp<sup>WT/F</sup>* control (**Figura 8 y Figura 9**).

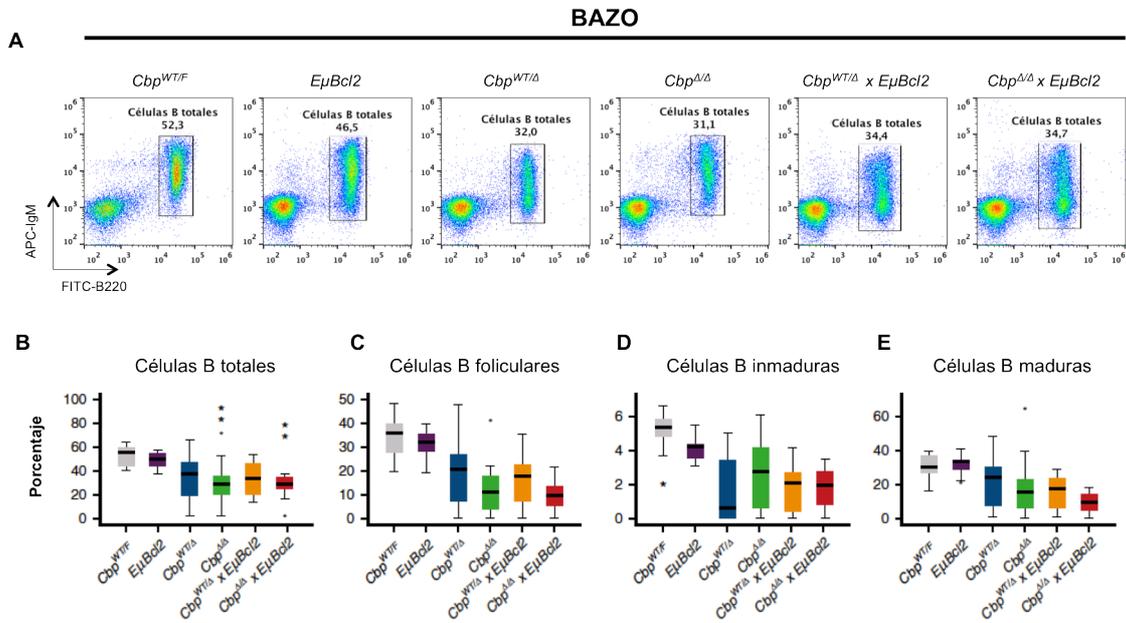
Curiosamente, aunque la sobreexpresión de *Bcl2* no fue capaz de rescatar completamente este déficit en el desarrollo B, observamos una tendencia hacia el incremento del número total de células B y del número de muchas subpoblaciones de células B en la médula ósea de los ratones *Cbp<sup>WT/Δ</sup> x EuBcl2* y *Cbp<sup>Δ/Δ</sup> x EuBcl2* en comparación con los ratones *Cbp<sup>WT/Δ</sup>* y *Cbp<sup>Δ/Δ</sup>*, respectivamente (**Figura 8A-E**). Sin embargo, estas diferencias no fueron estadísticamente significativas ( $P > 0,05$ ). Debido a la reducida magnitud de las diferencias observadas en el comportamiento de las células pre o pro-B con delección de *Crebbp* en comparación con las observadas en etapas posteriores de la diferenciación, sugerimos que la delección de *Crebbp* podría

estar afectando al desarrollo B, desde el estadio de célula pre o pro-B hasta el estadio en el que se empieza a expresar la inmunoglobulina M (IgM<sup>+</sup>), que corresponde a las células B inmaduras. Sin embargo, la etapa en la que la diferenciación de células B se ve perturbada no es claramente identificable en nuestros datos.



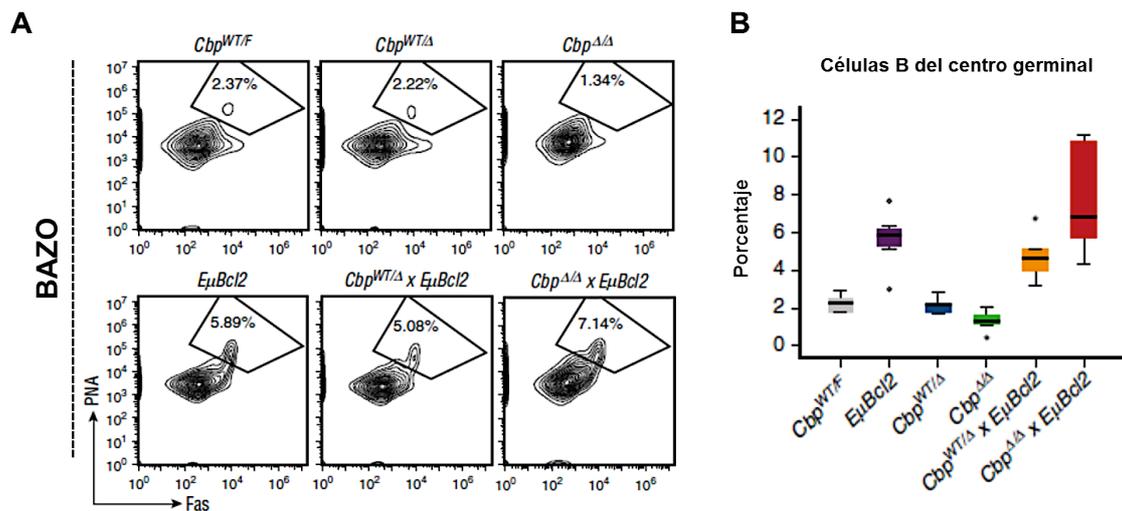
**Figura 8: La delección de *Crebbp* conduce a alteraciones en el desarrollo de células B de la médula ósea.** **A)** Ejemplos de análisis de citometría de flujo que muestran la tinción B220 e IgM de células de médula ósea de ratones mayores de 5 meses. Se puede observar un cambio notable en las frecuencias de las poblaciones de células B B220<sup>+</sup> entre cepas, estando asociada su reducción a la delección de *Crebbp* y siendo parcialmente rescatadas por el transgén de *EμBcl2*. **B)** Se muestra un resumen del porcentaje total de células B en la médula ósea, medido como el porcentaje de células B220<sup>+</sup>, para todas las cepas. El análisis mediante ANOVA unidireccional mostró una varianza significativa entre las cepas en el conjunto de los datos ( $P < 0.001$ ). La prueba post ANOVA (test de Tukey) reveló que esto fue impulsado por números significativamente mayores de células B totales en ratones *EμBcl2* en comparación con *Cbp*<sup>WT/F</sup> ( $P = 0.011$ ), *Cbp*<sup>WT/Δ</sup> ( $P < 0.001$ ), *Cbp*<sup>Δ/Δ</sup> ( $P < 0.001$ ) y *Cbp*<sup>WT/Δ</sup> x *EμBcl2* ( $P < 0.001$ ), pero no entre los ratones *EμBcl2* y *Cbp*<sup>Δ/Δ</sup> x *EμBcl2* ( $P = 0.068$ ). El resto de comparaciones no fueron significativas ( $P > 0.05$ ). **C)** Resumen del porcentaje de células pre o pro-B de médula ósea, medido como el porcentaje de células B220<sup>low</sup> IgM<sup>+</sup>, para todas las cepas de ratones mayores de 5 meses. No hubo varianza significativa para esta población entre las cepas (ANOVA unidireccional,  $P = 0.652$ ). **D)** Resumen de las frecuencias de células B inmaduras en médula ósea para todas las cepas de ratones mayores de 5 meses, medida como el porcentaje de células B220<sup>low</sup> IgM<sup>+</sup>. El análisis ANOVA unidireccional mostró una varianza significativa a través del conjunto de los datos ( $P = 0.034$ ) que fue impulsada por frecuencias significativamente más altas de células B inmaduras en la cepa *Cbp*<sup>Δ/Δ</sup> x *EμBcl2* comparadas con las de la cepa de *Cbp*<sup>WT/Δ</sup> (test de Tukey,  $P = 0.018$ ). No hubo otras comparaciones estadísticamente significativas en la prueba post ANOVA-test de Tukey ( $P > 0.05$ ). **E)** Por último, se muestra un resumen de las frecuencias de células B recirculantes para todas las cepas de ratones mayores de 5 meses, medidas como el porcentaje de células B220<sup>+</sup> IgM<sup>+</sup>. El análisis ANOVA unidireccional mostró una varianza significativa a través de las cepas en el conjunto de los datos ( $P < 0.001$ ). La prueba post ANOVA (test de Tukey) reveló que esto estaba impulsado por una frecuencia significativamente mayor de células B recirculantes en ratones *EμBcl2* en comparación con la de los ratones *Cbp*<sup>WT/F</sup> ( $P = 0.001$ ), *Cbp*<sup>WT/Δ</sup> ( $P < 0.001$ ), *Cbp*<sup>Δ/Δ</sup> ( $P < 0.001$ ), *Cbp*<sup>WT/Δ</sup> x *EμBcl2* ( $P < 0.001$ ) y *Cbp*<sup>Δ/Δ</sup> x *EμBcl2* ( $P = 0.004$ ).

Aunque también se detectaron reducciones en el porcentaje de células B en el bazo, no se observó el efecto de rescate del transgén *EμBcl2* en la frecuencia de células B, que sí vimos en la médula ósea, lo que sugiere que esto podría ser un efecto exclusivo de la sobreexpresión de *Bcl2* sobre el desarrollo de células B de la médula ósea, no alterando la supervivencia de las células B de la periferia (**Figura 9A-E**).



**Figura 9: La delección de *Crebbp* conduce a alteraciones en el desarrollo de células B del bazo. A)** Ejemplos de análisis de citometría de flujo que muestran la tinción B220 e IgM de esplenocitos de ratones mayores de 5 meses. Se puede observar una reducción en la frecuencia de células B totales ( $B220^+ IgM^{+/-}$ ) de los ratones con delección de *Crebbp*, y esto no se restaura mediante la adición del transgén *EμBcl2*. **B)** El gráfico de cajas muestra el porcentaje de células  $B220^+$  totales en el bazo de las 6 cepas de ratón. Hubo una variabilidad significativa entre las cepas (ANOVA unidireccional,  $P=0.007$ ) impulsado por frecuencias significativamente más bajas de células B totales en la cepa  $Cbp^{WT/\Delta}$  en comparación con las de las cepas  $Cbp^{WT/F}$  y  $Cbp^{WT/\Delta} \times E\mu Bcl2$  (test de Tukey,  $P=0.036$  y  $P=0.048$ , respectivamente). Ninguna otra comparación directa fue significativa. **C)** La gráfica muestra el porcentaje de células B foliculares en el bazo ( $B220^+ CD21^+ CD23^+$ ), a través de las 6 cepas. Hubo una varianza significativa entre las cepas, impulsada por frecuencias significativamente más bajas de células B foliculares en las cepas de ratones  $Cbp^{WT/\Delta}$ ,  $Cbp^{\Delta/\Delta}$ ,  $Cbp^{WT/\Delta} \times E\mu Bcl2$  y  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$  en comparación tanto con la cepa  $Cbp^{WT/F}$  como con  $E\mu Bcl2$  (test de Tukey,  $P < 0,01$  para todas las comparaciones uno a uno). **D)** El porcentaje de células B inmaduras en el bazo ( $B220^+ IgM^+ IgD^-$ ), se muestra en el gráfico de cajas. Hubo una varianza significativa a través de las cepas (ANOVA unidireccional,  $P < 0.001$ ) que incluyó frecuencias significativamente más bajas de células B inmaduras en las cepas  $Cbp^{WT/\Delta}$  y  $Cbp^{\Delta/\Delta}$  (test de Tukey,  $P < 0.001$  y  $P = 0.005$ , respectivamente) en comparación con ratones  $Cbp^{WT/F}$ . Además, también se observaron frecuencias significativamente más bajas de esta población en ratones  $Cbp^{WT/\Delta} \times E\mu Bcl2$  y  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$  (test de Tukey,  $P = 0.006$  y  $P = 0.013$ , respectivamente) en comparación con los ratones  $E\mu Bcl2$ . **E)** El gráfico de cajas muestra el porcentaje de células B maduras del bazo ( $B220^+ IgM^+ IgD^+$ ), a través de las 6 cepas. Hubo una varianza significativa a través de las cepas (ANOVA unidireccional,  $P < 0,001$ ) que fue impulsada por frecuencias significativamente más bajas en ratones  $Cbp^{\Delta/\Delta}$ ,  $Cbp^{WT/\Delta} \times E\mu Bcl2$  y  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$  en comparación tanto con las de los ratones  $Cbp^{WT/F}$  como con las de los ratones  $E\mu Bcl2$  (Tukey,  $P < 0,05$  para todas las comparaciones). No hubo diferencias significativas para la frecuencia de células B maduras entre los ratones  $Cbp^{WT/\Delta}$  y los ratones  $Cbp^{WT/F}$  (test de Tukey,  $P = 0.172$ ) y  $E\mu Bcl2$  (test de Tukey,  $P = 0.082$ ).

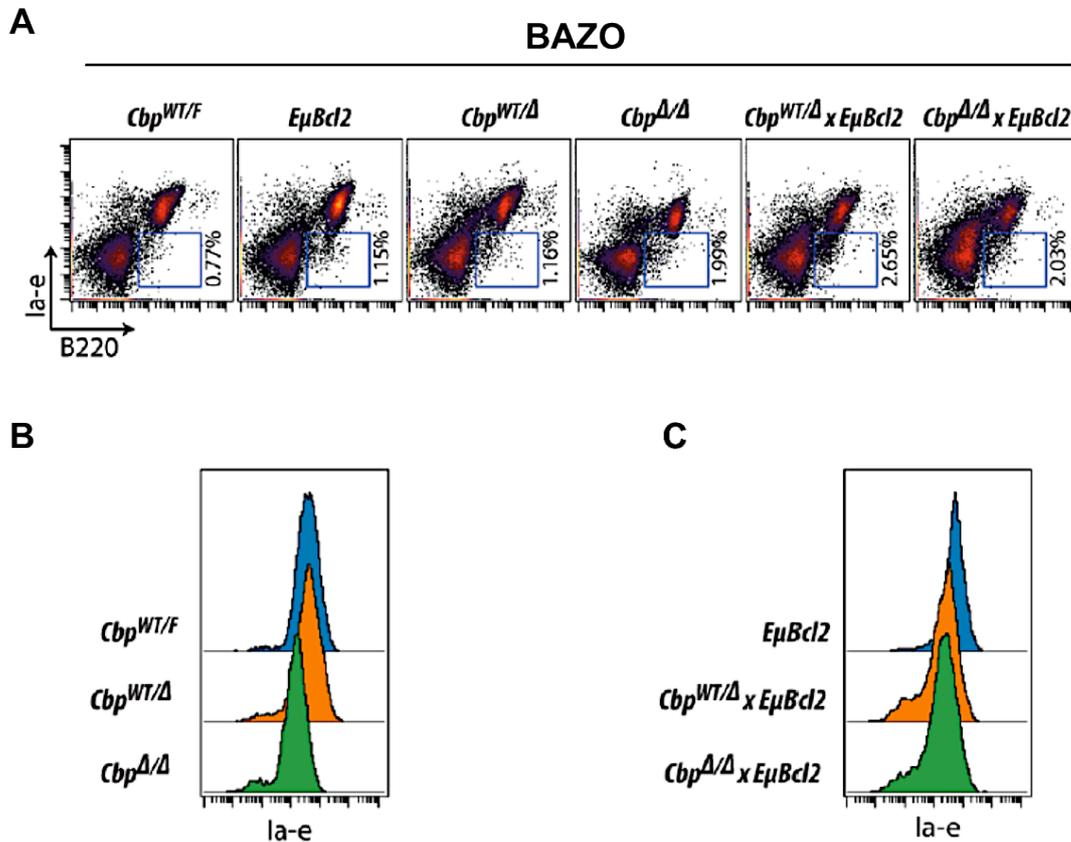
Por otro lado, observamos que los ratones  $Cbp^{WT/\Delta}$  y  $Cbp^{\Delta/\Delta}$  presentaban una capacidad reducida para producir células B del centro germinal en comparación con los ratones  $Cbp^{WT/F}$ , tras la inmunización de dichos animales con eritrocitos de cordero. Sin embargo, esta reducción no fue estadísticamente significativa. La generación de células B del centro germinal fue significativamente mayor en todos los ratones que portaban el transgén  $E\mu Bcl2$  ( $P < 0.001$ ) en comparación con las cepas sin el transgén; sin embargo, el desarrollo de células B del centro germinal no se vio afectado significativamente por la delección del gen *Crebbp* en este contexto (**Figura 10**).



**Figura 10: Estudio de la formación de células B del centro germinal.** Los ratones de cada cepa se inmunizaron mediante inyección intraperitoneal de  $1-2 \times 10^8$  glóbulos rojos de cordero, analizando mediante citometría de flujo el desarrollo de células B del centro germinal de los bazo, 10 días después de la inmunización. **A)** Imágenes de citometría de flujo representativas de las frecuencias de células B del centro germinal de las 6 cepas en estudio. Las frecuencias finales se obtuvieron a partir de las células viables (yoduro de propidio-negativas), B220-positivas. Las células B del centro germinal se definen como aglutinina de cacahuete-positivas (PNA) y Fas (CD95)-positivas. **B)** El gráfico de cajas muestra la tendencia de formación de células B del centro germinal de las diferentes cepas de ratones en estudio. En total se analizaron 8 ratones  $Cbp^{WT/F}$ , 8 ratones  $E\mu Bcl2$ , 5 ratones  $Cbp^{WT/\Delta}$ , 7 ratones  $Cbp^{\Delta/\Delta}$ , 6 ratones  $Cbp^{WT/\Delta} \times E\mu Bcl2$ , y 6 ratones  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$ . El análisis estadístico (ANOVA unidireccional) mostró una variabilidad significativa entre todas las cepas ( $P < 0.001$ ), pero el análisis post-anova (test de Tukey) demostró que esto venía dado por frecuencias significativamente más altas de células B de centro germinal en ratones que portaban el transgén  $E\mu Bcl2$  en comparación con ratones sin el transgén  $E\mu Bcl2$  ( $P < 0.001$  a  $0.040$ ). No hubo diferencias significativas entre las cepas  $Cbp^{WT/F}$  y  $Cbp^{WT/\Delta}$  o  $Cbp^{\Delta/\Delta}$ , ni entre las cepas  $E\mu Bcl2$  y  $Cbp^{WT/\Delta} \times E\mu Bcl2$  o  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$  ( $P > 0.005$ ).

Se ha observado que los pacientes con linfoma folicular primario y mutación de *CREBBP* presentan una reducción significativa en la expresión del complejo mayor de histocompatibilidad de clase II (MHC-II)<sup>143</sup>. Por ello, decidimos determinar si esto también estaba ocurriendo en los ratones con *Crebbp* inactivado. Con este fin, medimos mediante citometría de flujo los niveles de expresión de MHC-II en los bazo de todas las cepas de ratón en estudio (**Figura 11A**). Se observaron cambios menores en la expresión de MHC-II en aquellos ratones con *Crebbp* delecionado (**Figura 11B-C**), similares a los observados en el ratón *knock-down* de *Crebbp* generado mediante

ARN de interferencia (ARNi)<sup>291</sup>. Sin embargo, estas diferencias estaban muy por debajo de la magnitud de los cambios observados en los linfomas foliculares humanos<sup>143</sup>.

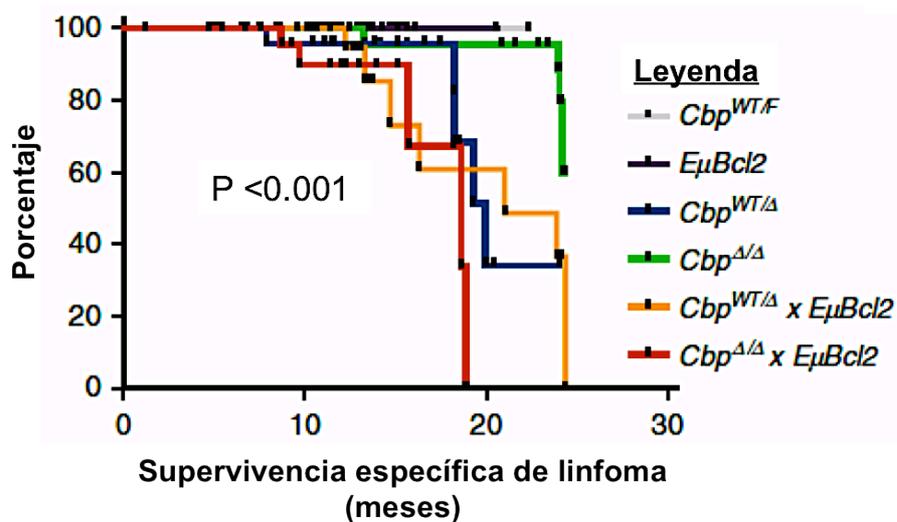


**Figura 11: Estudio de la expresión del complejo superior de histocompatibilidad de clase II (MHC-II) en los bazo de las cepas en estudio.** Mediante análisis de citometría de flujo se observaron pequeños cambios en la expresión de MHC-II como consecuencia de la delección de *Crebbp*. **A)** Ejemplos representativos de la expresión de MHC-II (Ia-e) en células B de bazo de las diferentes cepas de ratón. **B)** Histogramas representativos de la expresión de Ia-e en células B220+ de ratones sin el transgén *EμBcl2*. Se observa una modesta reducción de la expresión de MHC-II en las células B de los ratones *Cbp<sup>Δ/Δ</sup>* comparada con la de los ratones *Cbp<sup>WT/F</sup>*. **C)** Histogramas representativos de la expresión de Ia-e en células B220+ de ratones con el transgén *EμBcl2*. Se observa una modesta reducción de la expresión de MHC-II en las células B de los ratones *Cbp<sup>WT/Δ</sup> x EμBcl2* y *Cbp<sup>Δ/Δ</sup> x EμBcl2* comparada con la de los ratones *EμBcl2*.

En resumen, todos estos datos muestran que la inactivación de *Crebbp*, a nivel de células pro-B, produce una perturbación en el desarrollo normal de las células B, el cual se restablece parcialmente, pero no significativamente, mediante la sobreexpresión de *Bcl2*.

**1.2 La pérdida de *Crebbp* combinada con la sobreexpresión de *Bcl2* en células pro-B conduce al desarrollo de linfomas de células B.**

Para valorar el papel de *Crebbp* en el desarrollo de linfomas, monitorizamos toda la cohorte de ratones hasta que detectamos que los animales presentaban signos de enfermedad. Así, no observamos desarrollo de linfoma en las cepas *Cbp<sup>WT/F</sup>* ni *EμBcl2*. Sin embargo, los ratones de cepas con delección de *Crebbp* (*Cbp<sup>WT/Δ</sup>*, *Cbp<sup>Δ/Δ</sup>*, *Cbp<sup>WT/Δ</sup>* x *EμBcl2* y *Cbp<sup>Δ/Δ</sup>* x *EμBcl2*) enfermaron, mostrando pérdida de peso y menor actividad. En concreto, se observó una baja penetrancia y una larga latencia de la enfermedad en los ratones *Cbp<sup>WT/Δ</sup>* mientras que los ratones *Cbp<sup>Δ/Δ</sup>* mostraron una mayor penetrancia y una disminución de la latencia de la enfermedad. Sin embargo, la supervivencia específica de linfoma tanto de los ratones *Cbp<sup>WT/Δ</sup>* como de los ratones *Cbp<sup>Δ/Δ</sup>* no fue significativamente diferente de la de los ratones control *Cbp<sup>WT/F</sup>* (test Mantel-Cox *log-rank*, *P* =0.460 y 0.069, respectivamente) (**Figura 12, Tabla III**). Por otro lado, la penetrancia del linfoma se vio incrementada y la latencia disminuida, en cada una de estas cepas de ratón, con la adición del transgén *EμBcl2*. Así, los ratones *Cbp<sup>WT/Δ</sup>* x *EμBcl2* y *Cbp<sup>Δ/Δ</sup>* x *EμBcl2* presentaron una supervivencia específica de linfoma significativamente menor en comparación con la de los ratones control que únicamente portaban el transgén *EμBcl2* (test Mantel-Cox *log-rank*, *P* =0.048 y 0.025, respectivamente) (**Figura 12, Tabla III**). Aunque no hay ninguna evidencia que sugiera que el transgén *mb1-cre* promueve el linfoma, tras ser usado por múltiples laboratorios en todo el mundo, quisiéramos reseñar que no incluimos en el estudio ratones control sólo con el transgén *mb1-cre* y, por lo tanto, no podemos excluir una posible contribución a la linfomagénesis del mismo.



**Figura 12: Supervivencia específica para el desarrollo de linfoma de células B.** La gráfica Kaplan-Meier muestra la supervivencia específica de linfoma de las 6 cepas de ratones transgénicos. Como se muestra, la delección de uno o ambos alelos de *Crebbp* conduce al desarrollo de linfoma a edades avanzadas, sin embargo la penetrancia se ve incrementada y la latencia disminuida con la adición del transgén *EμBcl2*.

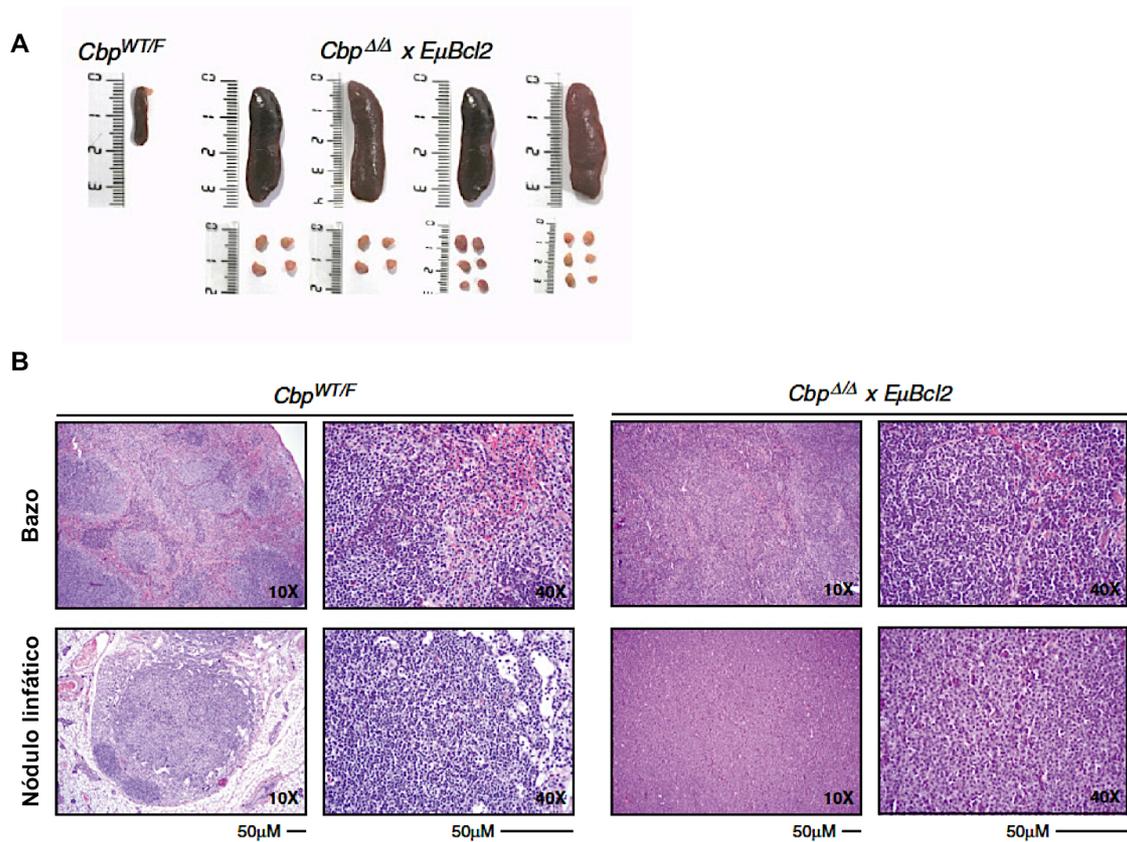
Tabla III: Resumen de la penetrancia y latencia media de linfoma de células B en las diferentes cepas transgénicas de ratón en estudio.

Cepa	N	Penetrancia media (%)	Latencia media (meses)
<i>Cbp</i> <sup>WT/F</sup>	16	0	NA
<i>EμBcl2</i>	24	0	NA
<i>Cbp</i> <sup>WT/Δ</sup>	37	10,8	21,37
<i>Cbp</i> <sup>Δ/Δ</sup>	27	18,5	16,73
<i>Cbp</i> <sup>WT/Δ</sup> x <i>EμBcl2</i>	33	21,2	17,99
<i>Cbp</i> <sup>Δ/Δ</sup> x <i>EμBcl2</i>	29	17,2	14,33

NA: no aplica

La necropsia de los ratones enfermos reveló esplenomegalia y agrandamiento de los ganglios linfáticos (**Figura 13A**), con infiltración linfocítica en los pulmones, hígado y riñones en algunos de los ratones. Los estudios de la histología tumoral, llevados a cabo por hematopatólogos experimentados de la Universidad de Nebraska (USA), mostraron que la mayoría de estos tumores eran más similares, desde un punto de vista histológico, a los linfomas foliculares de grado 3 y a los DLBCL (**Figura 13B**). Un subconjunto de tumores de la cepa *Cbp*<sup>WT/Δ</sup> fue similar a los linfomas foliculares de grado 1 y 2, desde un punto de vista histopatológico.

Los estudios de citometría de flujo de los diferentes compartimentos celulares nos permitieron confirmar que estos ratones con delección de uno o dos alelos de *Crebbp* desarrollaban linfomas de células B. Así, encontramos que estos ratones desarrollaban con frecuencia DLBCL con inmunofenotipo de células B del centro germinal (**Figura 14**). Aunque con menor frecuencia, algunos ratones también desarrollaron DLBCL con inmunofenotipo de células B maduras (B220<sup>+</sup> IgM<sup>+</sup> IgD<sup>+</sup>) y de células B inmaduras (B220<sup>+</sup> IgM<sup>+</sup> IgD<sup>-</sup>). Desde un punto de vista inmunofenotípico, no detectamos ningún caso de linfoma folicular.



**Figura 13: La delección de *Crebbp* en tándem con la sobreexpresión de *Bcl2* promueve el desarrollo de linfomas de células B agresivos. A)** Ejemplo representativo del bazo y los nódulos linfáticos de ratones *Cbp<sup>Δ/Δ</sup> x EμBcl2* enfermos. En la figura se muestra como referencia el bazo de un ratón control *Cbp<sup>WT/F</sup>* de edad similar. Como se observa, los ratones enfermos desarrollaron esplenomegalia y linfadenopatías, ambas características del linfoma humano. **B)** Tinción hematoxilina-eosina de un bazo y un nódulo linfático representativos de los ratones que desarrollaron linfomas. En la figura se puede observar el corte histológico de un bazo y un nódulo linfático de un ratón *Cbp<sup>Δ/Δ</sup> x EμBcl2* enfermo, comparado con el bazo y el nódulo linfático de un ratón control *Cbp<sup>WT/F</sup>* de la misma edad. La revisión llevada a cabo por los hematopatólogos determinó que este espécimen era más similar a un DLBCL desde un punto de vista histológico.

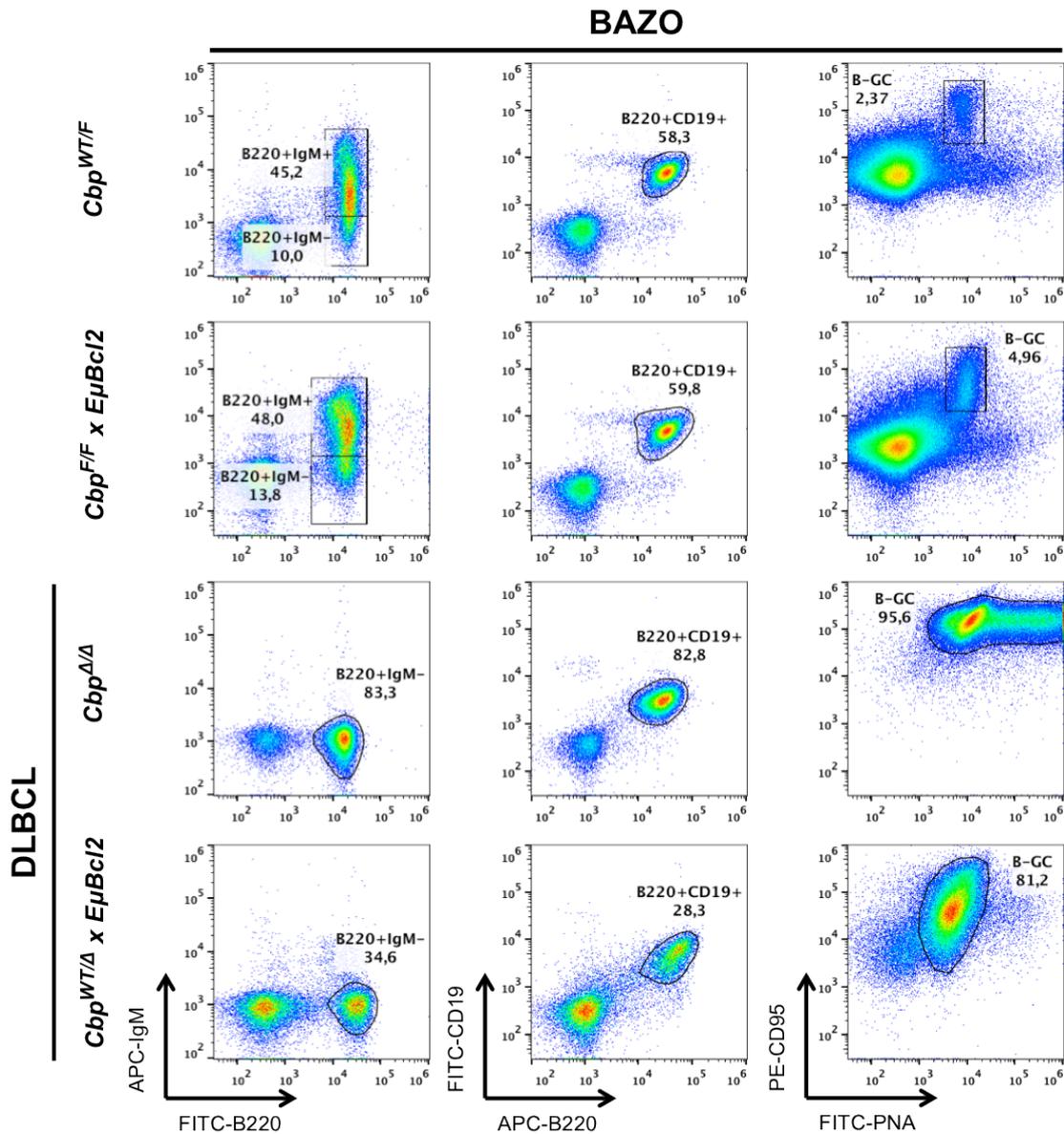
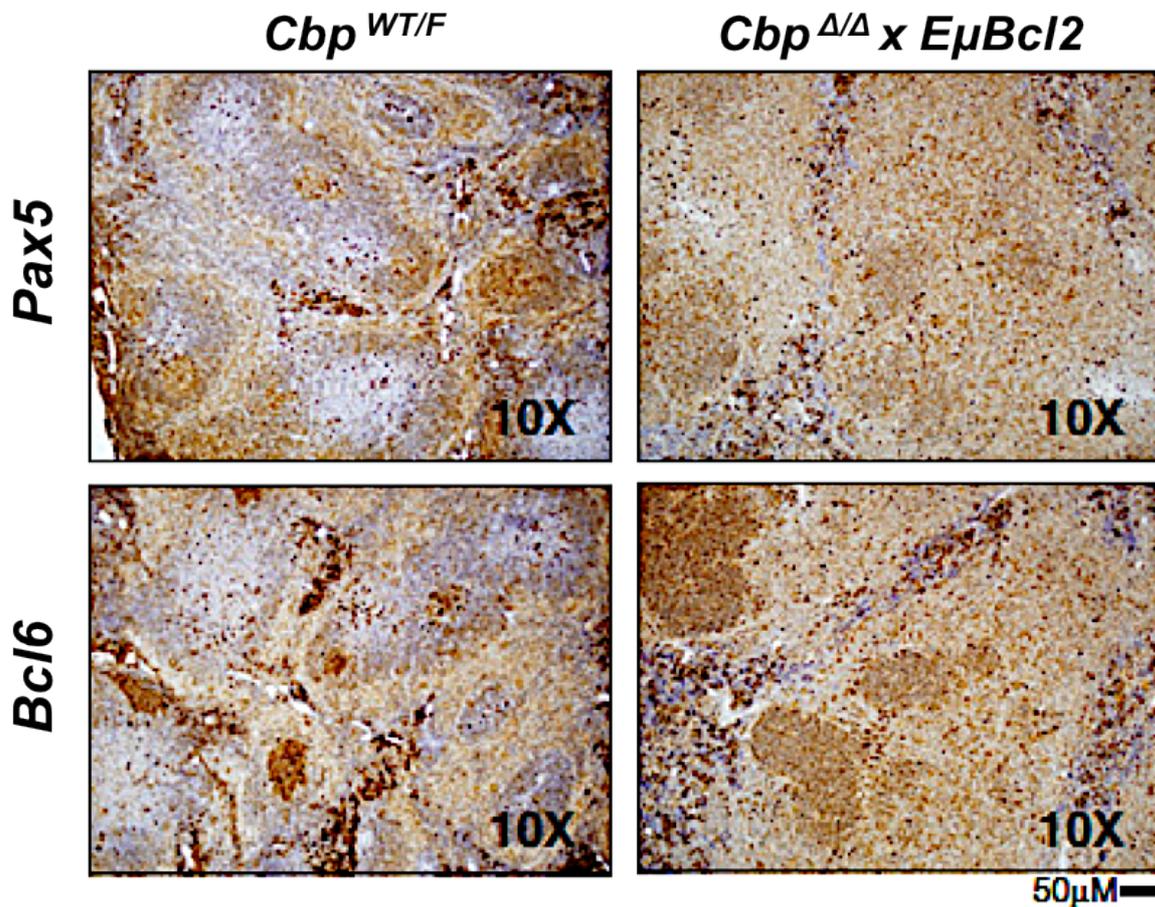


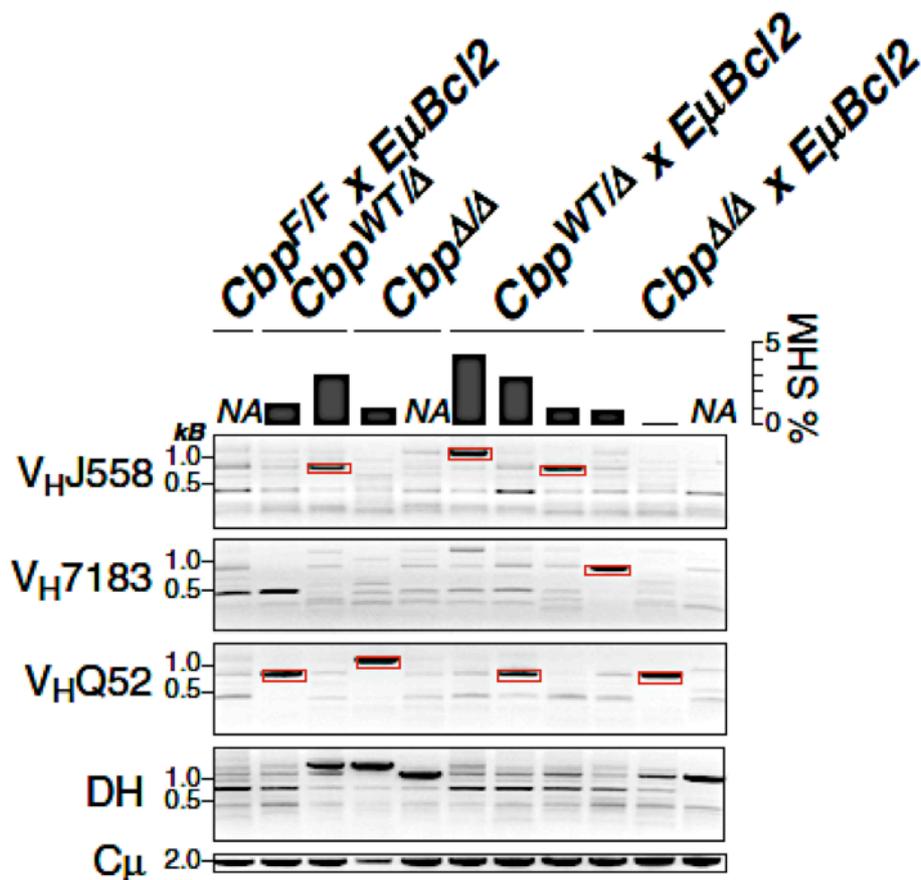
Figura 14: Los ratones con delección de uno o dos alelos de *Crebbp* desarrollan con frecuencia linfomas de células B (DLBCL) con un inmunofenotipo característico de células B del centro germinal. Se muestran dos ejemplos representativos del inmunofenotipo de los linfomas que con mayor frecuencia desarrollaron los ratones con delección de *Crebbp*. La población de células B de centro germinal se define como B220<sup>+</sup>CD95<sup>+</sup>PNA<sup>+</sup>. Los bazo de un ratón *Cbp*<sup>WT/F</sup> y de un ratón *EμBcl2* de edad similar se muestran como referencia (estos ratones control fueron inmunizados para observar la reacción de centro germinal no patológica).

Los estudios de inmunohistoquímica mostraron que todos los tumores eran uniformemente positivos para Pax5, lo que indica que estos tumores tenían una identidad celular B. Además, dichos estudios también nos permitieron determinar que todos los tumores eran positivos para Bcl6, confirmando que se trataba de tumores de células B de centro germinal (**Figura 15**).



**Figura 15: Los linfomas de ratones *Crebbp* transgénicos tienen características de células B del centro germinal.** En la figura se muestra un ejemplo representativo de tinción inmunohistoquímica para Pax5 y Bcl6 en el bazo de un ratón *Cbp<sup>Δ/Δ</sup> x EμBcl2* enfermo comparado con el bazo de un ratón control *Cbp<sup>WT/F</sup>* de la misma edad. El bazo del ratón control muestra folículos benignos normales y el patrón de tinción para Pax5 y Bcl6 esperado. El bazo del ratón *Cbp<sup>Δ/Δ</sup> x EμBcl2* con linfoma muestra una histología difusa con células que son Pax5 y Bcl6 positivas, lo que apoya que estos linfomas son característicos de células B de centro germinal.

Se llevó a cabo el análisis de los reordenamientos VDJ de alguno de los tumores, mediante reacción en cadena de la polimerasa (PCR), y se observó que la mayoría de los tumores poseían un reordenamiento clonal VDJ con marcas de hipermutación somática (**Figura 16**). Estos datos hacen aún más consistente la idea de que estos tumores tienen características de células B de centro germinal (GCB).



**Figura 16: Los linfomas procedentes de ratones con *Crebbp* deletado muestran un origen clonal.** Los reordenamientos de las inmunoglobulinas se evaluaron mediante PCR del ADN extraído de bazo implicados en linfomas (n=10) y un bazo de ratón control de edad similar (*Cbp<sup>F/F</sup> x EμBcl2*). El ratón control mostró un patrón policlonal indicativo de una población de células B normal, como se muestra en el carril más a la izquierda. Por el contrario, la mayoría de las muestras tumorales de los ratones *Cbp<sup>WT/Δ</sup>*, *Cbp<sup>Δ/Δ</sup>*, *Cbp<sup>WT/Δ</sup> x EμBcl2* y *Cbp<sup>Δ/Δ</sup> x EμBcl2* mostraron una única banda dominante que es indicativa de una población de células B clonales, tal y como se observa en los ejemplos de la figura. Ocho de estas bandas, destacadas con rectángulos rojos, se escindieron y se clonaron para su secuenciación. El análisis de secuenciación reveló la presencia de hipermutación somática en 7/8 tumores, con una desviación media del 2% de la secuencia del gen V de la línea germinal (min = 0,8%, max = 4,2%). Esto proporciona una evidencia adicional que apoyaría la idea de que estos tumores son característicos de células GCB, o de que las células B han transitado previamente a través del centro germinal. NA, no aplicable; SHM, hipermutación somática.

Como estos linfomas eran clonales y mostraban una larga latencia, nos planteamos rastrear posibles eventos genéticos secundarios que aparecieran durante el periodo de latencia de la enfermedad. Para ello, llevamos a cabo experimentos de secuenciación del exoma completo (WES por sus siglas en inglés, *Whole Exome Sequencing*) de 2 tumores de la cepa *Cbp<sup>WT/Δ</sup>* y de 4 tumores de la cepa *Cbp<sup>Δ/Δ</sup> x EμBcl2*, utilizando una profundidad media de 98X (min = 46X, max = 158X). Estos tumores poseían un promedio de 148 variantes somáticas codificantes (min = 34, max = 275), pero ninguno de los genes afectados ha sido previamente asociado al linfoma humano de células B (**Tabla IV**).

Tabla IV: Resumen de la secuenciación del exoma completo de los linfomas de células B desarrollados por los ratones *Cbp<sup>WT/Δ</sup>* y *Cbp<sup>Δ/Δ</sup> x EμBcl2*.

Cepa	ID	% acierto	Media de lecturas	Mutaciones codificantes
Control	SL209716	84,15%	135.505782	NA
<i>Cbp<sup>WT/Δ</sup></i>	SL209715	85,54%	101.638036	122
<i>Cbp<sup>WT/Δ</sup></i>	SL209714	84,38%	115.878537	275
<i>Cbp<sup>Δ/Δ</sup> x EμBcl2</i>	SL209719	88,99%	63.136878	34
<i>Cbp<sup>Δ/Δ</sup> x EμBcl2</i>	SL209720	81,78%	66.295523	75
<i>Cbp<sup>Δ/Δ</sup> x EμBcl2</i>	SL209718	65,41%	46.440094	186
<i>Cbp<sup>Δ/Δ</sup> x EμBcl2</i>	SL209717	83,73%	158.172422	197

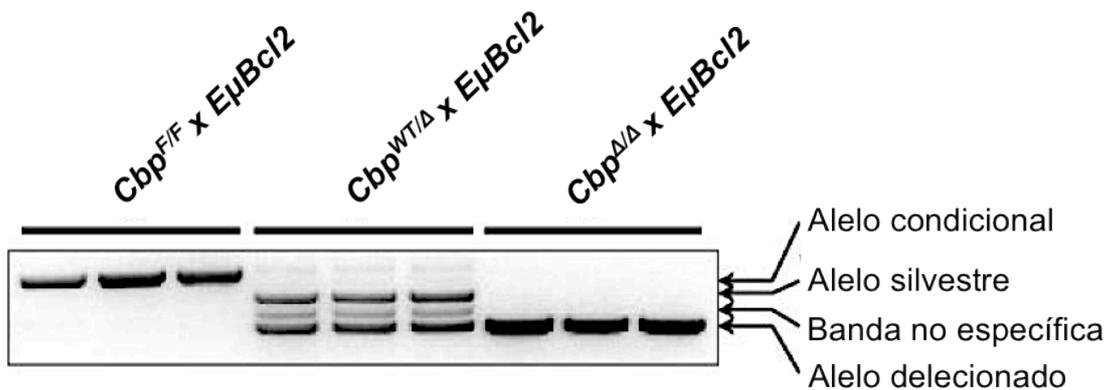
NA: no aplica

Todos estos datos demuestran que la inactivación de *Crebbp* induce un proceso de reprogramación tumoral a nivel de una célula pro-B que puede conducir al desarrollo espontáneo de DLBCL de células B de centro germinal, y este hecho se ve potenciado por la sobreexpresión de *Bcl2*.

### 1.3 La inactivación de *Crebbp* en células B induce de forma indirecta la expresión de *Myc*.

Con el objetivo de investigar los mecanismos tempranos que ocurren en la patogénesis del linfoma de células B asociado con la inactivación de *Crebbp*, aislamos células B de bazo procedentes de ratones adultos (de unos 10 meses de edad), libres de enfermedad, de las cepas *Cbp<sup>FF</sup> x EμBcl2* (n=3), *Cbp<sup>WT/Δ</sup> x EμBcl2* (n=3) y *Cbp<sup>Δ/Δ</sup> x EμBcl2* (n=3) para determinar su perfil transcripcional y epigenético. En primer lugar, quisimos confirmar la pureza de las células B aisladas, mediante PCR. De esta manera, confirmamos no sólo la alta pureza de las células B separadas, sino también, que la Cre-recombinasa estaba funcionando correctamente en las células B

de las cepas  $Cbp^{WT/\Delta} \times E\mu Bcl2$  y  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$ , puesto que se detectó el alelo(s) delecionado(s) de *Crebbp* (Figura 17).

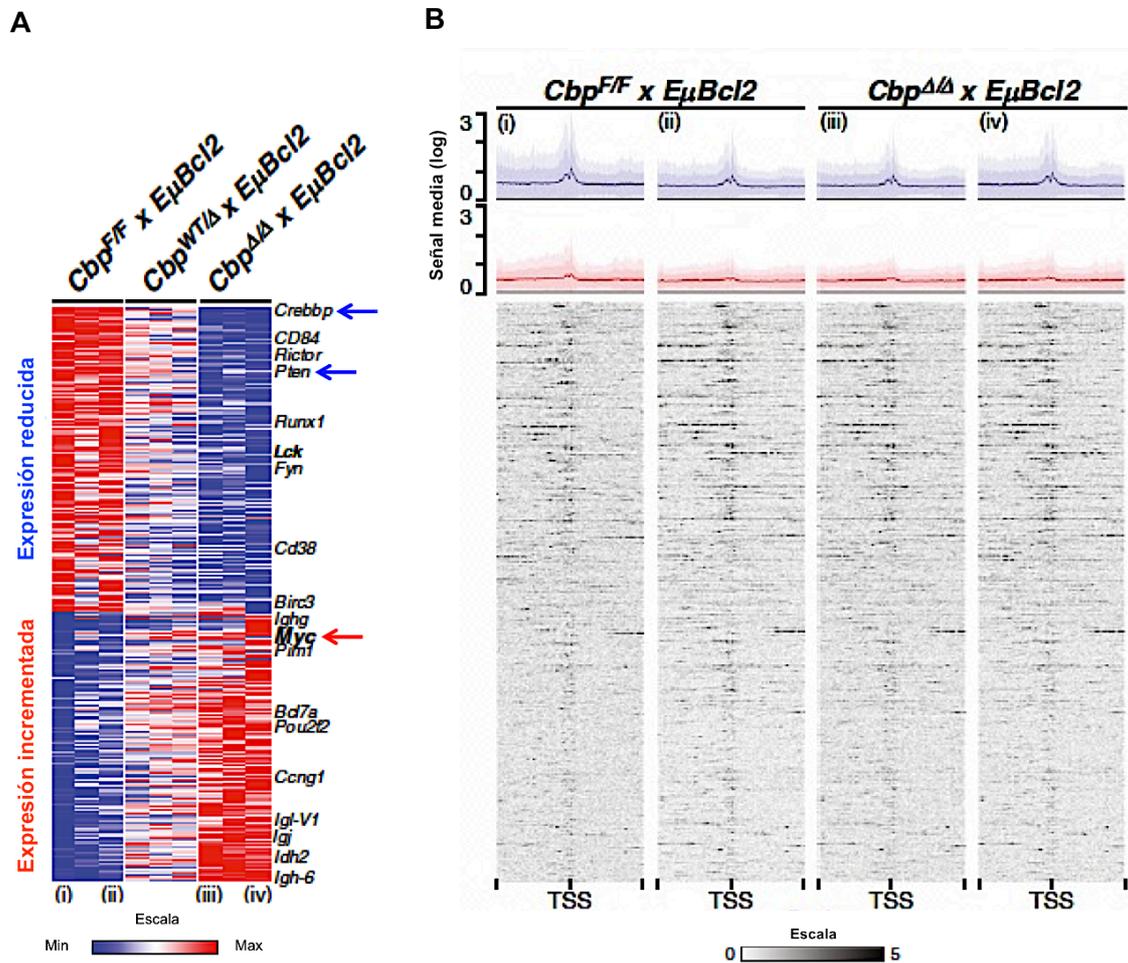


**Figura 17: Deleción del transgén *Crebbp* condicional en células B purificadas.** Análisis por PCR del ADN extraído de células B purificadas de bazo de las 3 cepas transgénicas que se indican. Se utilizó como control células B purificadas del bazo de ratones  $Cbp^{F/F} \times E\mu Bcl2$ , con el objetivo de detectar el alelo condicional de *Crebbp*. Si la pureza de las células no fuera alta, el alelo condicional debería ser visible en las muestras de los ratones  $Cbp^{WT/\Delta} \times E\mu Bcl2$  y  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$ , debido a la contaminación por células no B. Si la deleción del transgén condicional dirigida por *mb1-cre* no fuera eficiente, el alelo condicional debería ser visible en las muestras de los ratones  $Cbp^{WT/\Delta} \times E\mu Bcl2$  y  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$ , dado que el alelo condicional no se eliminaría. La ausencia del alelo condicional del transgén *Crebbp* en las muestras de los ratones  $Cbp^{WT/\Delta} \times E\mu Bcl2$  y  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$  confirma, primero, una alta pureza de la población de células B purificadas, y segundo, que el alelo condicional es eficientemente delecionado en las células B.

El análisis de expresión génica diferencial identificó 370 genes con una expresión significativamente reducida y 335 genes con un aumento significativo de la expresión en las células B de ratones con uno o ambos alelos de *Crebbp* delecionados y sobreexpresión de *Bcl2* (cepas  $Cbp^{WT/\Delta} \times E\mu Bcl2$  y  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$ ) en comparación con aquellos sin eliminación de *Crebbp* pero con sobreexpresión de *Bcl2* (cepa  $Cbp^{F/F} \times E\mu Bcl2$ ). Como era de esperar, se detectó una expresión reducida de *Crebbp* en las células B de los ratones con deleción de *Crebbp*. También se observaron otras alteraciones que podrían estar asociadas a la patogénesis de estos tumores, como la reducción de la expresión del gen *Pten*, el aumento de la expresión de múltiples genes asociados a las inmunoglobulinas, y el aumento de la expresión de *Myc* (Figura 18A).

Debido a que existen asociaciones previas entre las mutaciones en *CREBBP* y una reducción de la acetilación de la histona H3 en la lisina 18 (H3K18Ac)<sup>132,181</sup>, las células B de dos de los ratones  $Cbp^{F/F} \times E\mu Bcl2$  y de dos de los ratones  $Cbp^{\Delta/\Delta} \times E\mu Bcl2$ , analizados mediante arrays de expresión (marcados en el array como i, ii, iii, iv), fueron también analizadas mediante secuenciación de inmunoprecipitación de la cromatina ó ChIP-seq, con el objetivo de estudiar su perfil génico para H3K18Ac. Sin

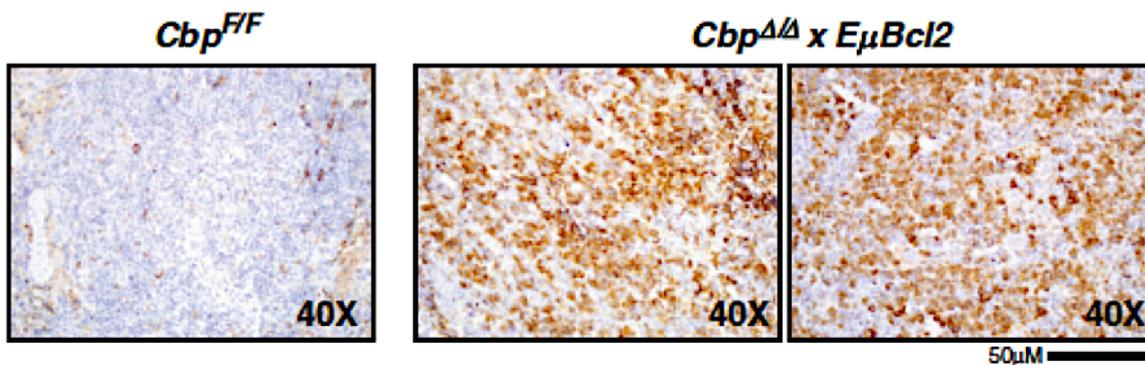
embargo, no se observaron diferencias significativas en los niveles de H3K18Ac (Figura 18B).



**Figura 18: Los cambios transcripcionales asociados con la delección de *Crebbp* no están asociados con cambios en la región proximal de H3K18Ac.** **A)** Perfil de expresión génica de células B purificadas del bazo de ratones adultos libres de enfermedad. Se observan alteraciones moleculares tempranas asociadas con la etiología de la enfermedad. El análisis de expresión génica diferencial reveló un grupo de genes con un significativo aumento o reducción de su expresión (FDR <0,25, fold-change >1,2) asociado con la delección de uno o ambos alelos de *Crebbp* dentro del fondo genético *EμBcl2*. Dentro de este grupo de genes, se pudo detectar un incremento en la expresión del oncogén *Myc*. Datos completos disponibles en *Gene expresión Omnibus*, GSE85490. **B)** Los mapas de calor muestran el nivel de expresión de H3K18Ac desde 10 kb aguas arriba hasta 10 kb aguas abajo de los sitios de inicio de la transcripción (TSS) de los genes con una expresión diferencial asociada a la delección de *Crebbp*. Se realizó ChIP-seq de cuatro de las muestras que fueron usadas para estudiar el perfil de expresión génica (nombradas en ambos paneles como i, ii, iii y iv). El nivel de H3K18Ac de cada gen se alinea con los niveles de expresión del panel A, y la señal media (línea) ± la desviación estándar (región sombreada) se resume para los genes con una expresión génica reducida (azul) o aumentada (roja) en la parte superior del mapa de calor. Hay un pico de la expresión de H3K18Ac en los TSS de la mayoría de los genes con expresión diferencial, que muestra nuestra capacidad para detectar H3K18Ac en TSSs. Sin embargo, los cambios en la expresión génica entre las cepas no se asociaron con cambios en H3K18Ac en los TSS ± 10 kb.

Por tanto, estos datos indican que aunque la delección de *Crebbp* está asociada con alteraciones transcripcionales significativas, que probablemente sean importantes para la biología de la enfermedad, tales como la inducción de la expresión de *Myc*, éstas no se deben a los cambios directos en H3K18Ac de regiones reguladoras proximales de genes, por lo que podrían estar siendo dirigidas por elementos reguladores distantes o factores secundarios.

Este aumento en la expresión de *Myc* fue confirmado, además, mediante inmunohistoquímica de los bazos de ratones *Cbp<sup>ΔΔ</sup> x EμBcl2* libres de enfermedad. La inmunohistoquímica mostró una fuerte tinción para *Myc* en el bazo de los ratones *Cbp<sup>ΔΔ</sup> x EμBcl2*, que no se detectaba en el bazo de los ratones control *Cbp<sup>F/F</sup>* (**Figura 19**). Dado que *Myc* es un importante oncogén, nos planteamos que la sobreexpresión de *Myc* quizás era una huella molecular que probablemente contribuía a la biología de la enfermedad.

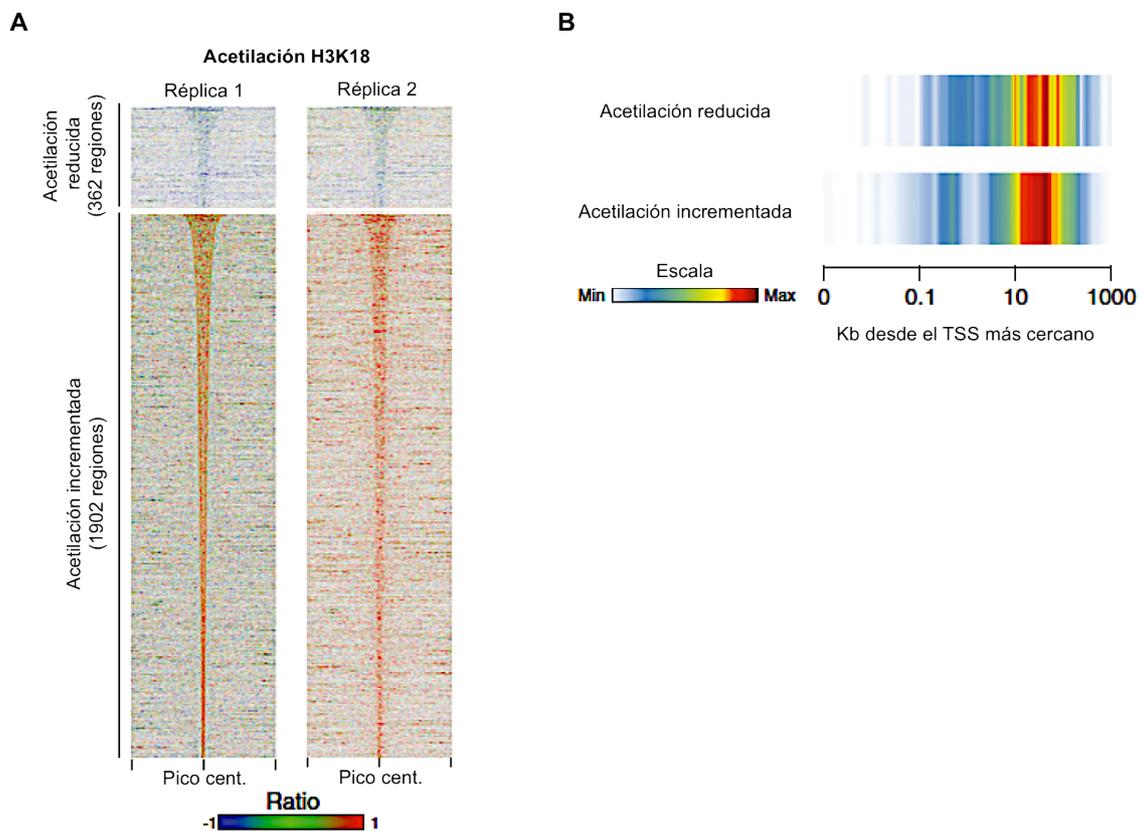


**Figura 19:** La sobreexpresión de *Myc* es una huella molecular asociada a la inactivación de *Crebbp*, que aparece en etapas pretumorales. El incremento de la expresión de *Myc* fue confirmado mediante tinción inmunohistoquímica en los bazos de ratones *Cbp<sup>Δ/Δ</sup> x EμBcl2* libres de enfermedad comparados con la de los controles *Cbp<sup>F/F</sup>* de la misma edad.

#### **1.4 La inactivación de *Crebbp* está relacionada con una firma epigenética asociada a *Myc*.**

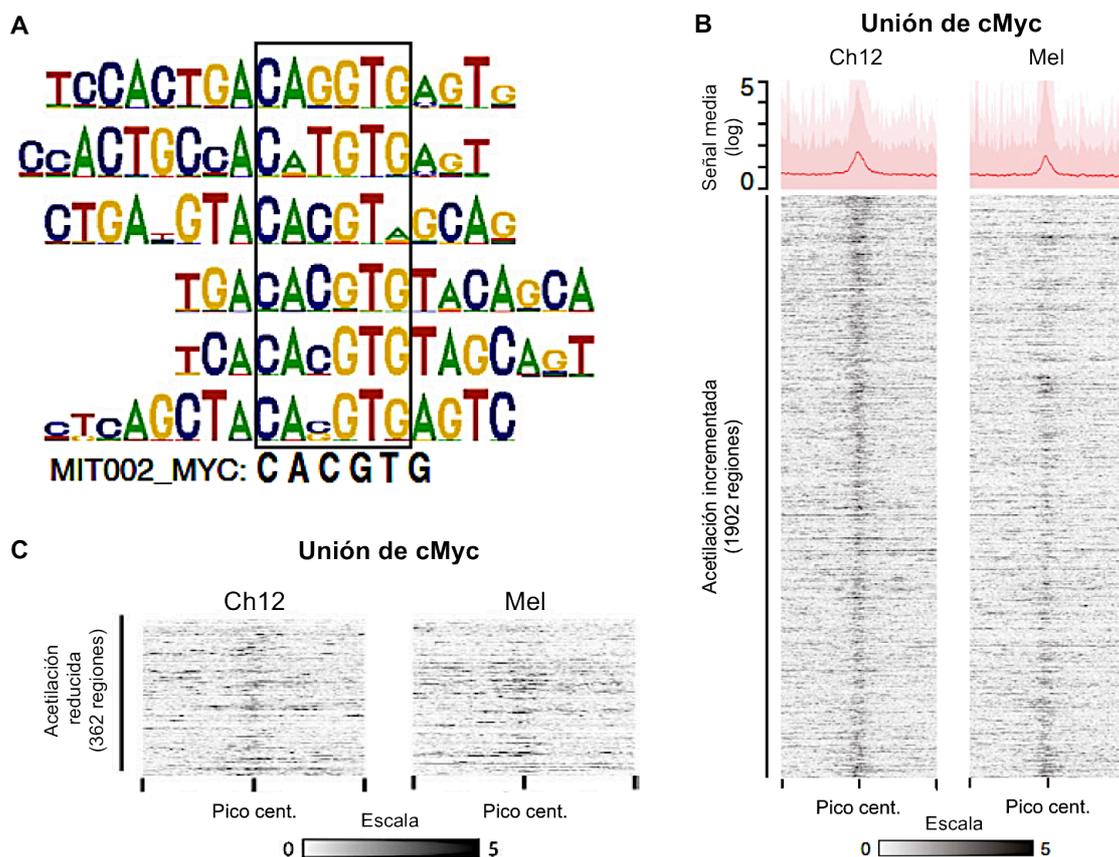
Aunque los cambios en H3K18Ac no se observaron en los TSSs (sitios de inicio de la transcripción) de genes expresados diferencialmente, sí que detectamos significativamente alterado H3K18Ac en otras regiones genómicas (**Figura 20A**). Específicamente, se identificaron 362 regiones con acetilación significativamente reducida (un total de 886 kb; kilobases) y 1902 regiones con acetilación significativamente incrementada (un total de 3487 kb) en células B de ratones *Cbp<sup>ΔΔ</sup> x EμBcl2* en comparación con los ratones control *Cbp<sup>F/F</sup> x EμBcl2*. Se trataba, predominantemente, de regiones intragénicas del genoma, con una distancia media de

68 kb o 77 kb al TSS más cercano para regiones de acetilación reducida o aumentada, respectivamente (**Figura 20B**). Esta localización es indicativa de los patrones alterados de regulación de H3K18Ac en regiones intragénicas potenciadoras. Para definir qué factores podían estar mediando o regulando estos cambios en H3K18Ac, se realizó un análisis de enriquecimiento de motivos del ADN. La comparación de secuencias de ADN de regiones con H3K18Ac significativamente alterado con regiones del mismo tamaño inmediatamente aguas arriba e inmediatamente aguas abajo, identificó 846 motivos enriquecidos que correspondían a 581 grupos homólogos de genes (*datos disponibles en supplemental materials for García-Ramírez et al; Blood 2017 129:2645-2656. Table S8*). Al comparar los grupos de motivos sobrerrepresentados con matrices de secuencias de unión al ADN de factores de transcripción conocidos, identificamos factores de transcripción que podrían unirse a estas regiones.



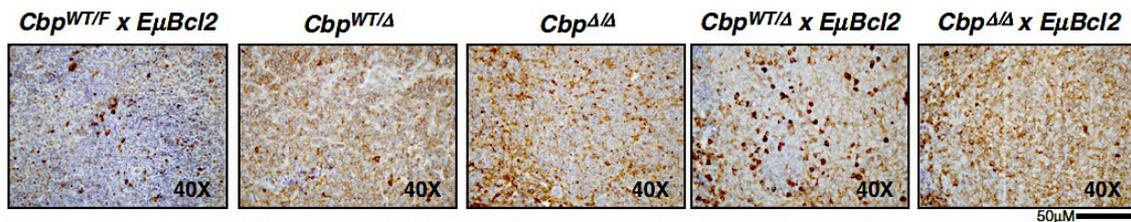
**Figura 20: Las regiones con diferente H3K18Ac son principalmente intragénicas. A)** Mapa de calor del nivel de H3K18Ac en células B de ratones  $Cbp^{\Delta/\Delta} \times EuBcl2$  comparadas con células B de ratones  $Cbp^{F/F} \times EuBcl2$  (los resultados son representativos de un experimento de 2 réplicas biológicas). Estos mapas muestran la intersección de las diferencias significativas identificadas en cada réplica, que incluye un pequeño número de regiones con picos significativamente reducidos de H3K18Ac y un gran número de regiones con un aumento significativo de picos de H3K18Ac. **B)** Las regiones de acetilación reducida y aumentada afectan principalmente a las regiones intragénicas que están distantes del TSS más cercano. La distancia desde el centro del pico de las regiones de acetilación significativamente reducida (arriba) o aumentada (abajo) al TSS del gen más próximo se muestra usando un diagrama de calor. Se puede ver que la mayoría de los picos con H3K18Ac diferencial están muy lejos del TSS más cercano, lo que sugiere que podrían ser elementos reguladores distantes. Cent., Centro.

En concordancia con la sobreexpresión observada de *Myc* en estas muestras, 14 de los grupos de motivos del ADN mostraban una homología significativa con el sitio de reconocimiento de *Myc* y estos grupos de motivos se encontraron en el 64% de las regiones con cambios en H3K18Ac (1454/2264) asociados con la inactivación de *Crebbp* (Figura 21A). Con el fin de proporcionar una evidencia adicional de que *Myc* se une a estas regiones con desregulación de H3K18Ac, se usaron datos públicos de *ChIP-seq* para *Myc* procedentes de líneas celulares murinas de linfoma de células B (Ch12) y de eritroleucemia (Mel). La señal *ChIP* para *Myc* mostró una fuerte unión al centro de los picos de las regiones con un aumento de la acetilación en células B de ratones *Cbp<sup>ΔΔ</sup> x EμBcl2* en comparación con los ratones *Cbp<sup>F/F</sup> x EμBcl2* (Figura 21B). La señal de *Myc* fue particularmente notable en la línea celular derivada de las células B, Ch12. Ciertas regiones con disminución de la acetilación también mostraron unión de *Myc* en células B de ratones *Cbp<sup>ΔΔ</sup> x EμBcl2* en comparación con ratones *Cbp<sup>F/F</sup> x EμBcl2* (Figura 21C).



**Figura 21: Las regiones con alteraciones de H3K18Ac están vinculadas con *Myc*.** **A)** Ejemplo de motivos más significativamente enriquecidos en regiones alteradas de H3K18Ac, que muestran homología con los sitios de unión de MYC. **B)** Datos de *ChIP-seq* para *Myc* en líneas celulares murinas de linfoma de células B (Ch12) y de eritroleucemia (Mel) muestran un pico de unión fuerte de *Myc* en la misma localización que el pico de acetilación aumentada de H3K18 observado en células B de ratones *Cbp<sup>ΔΔ</sup> x EμBcl2*. La señal media (línea) ± desviación estándar (región sombreada) se resume en la parte superior del mapa de calor. **C)** Análisis de *ChIP-seq* para *Myc* en células Ch12 y Mel a través de regiones con H3K18Ac reducida en células B de ratones *Cbp<sup>ΔΔ</sup> x EμBcl2* en comparación con células B de ratones *Cbp<sup>F/F</sup> x EμBcl2* ratones. La unión de *Myc* se puede observar en el centro del pico, aunque la intensidad de la señal es menor que la observada en las regiones con H3K18Ac incrementada. Cent= Centro; *cMyc*= *Myc* protooncogén.

Estos resultados indicaban que *Myc* podía estar desempeñando un importante papel en la etiología de la enfermedad. Por ello, nos planteamos confirmar la expresión de *Myc* mediante tinción inmunohistoquímica en un amplio número de bazos tumorales, lo que nos permitió observar una fuerte expresión de *Myc* en los bazos de todas las cepas con uno o ambos alelos de *Crebbp* deletado en comparación con la expresión observada en el bazo de ratones control *Cbp<sup>WT/F</sup> x EμBcl2* (Figura 22).



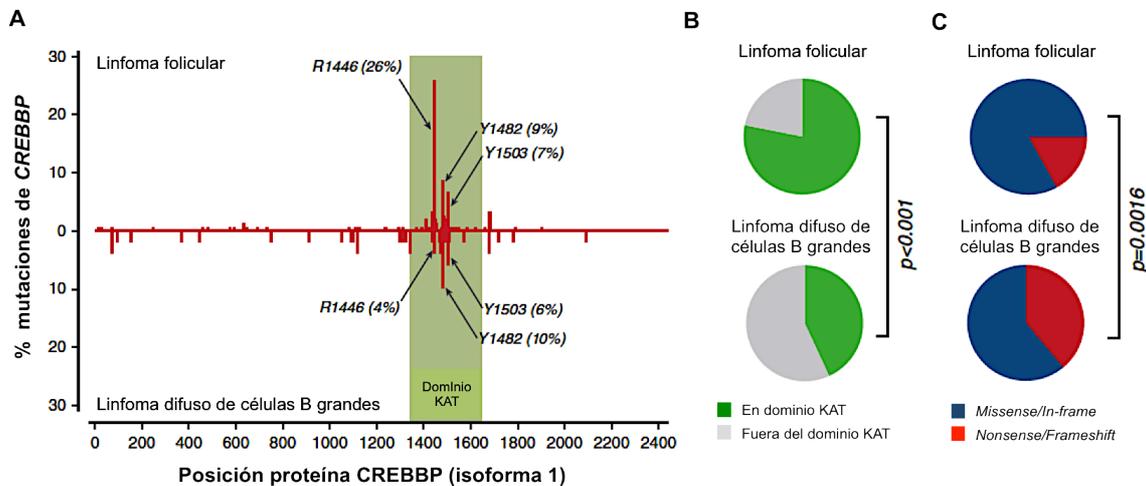
**Figura 22: Expresión de Myc en las células tumorales de los ratones transgénicos con delección de *Crebbp*.** Debido a la importancia potencial de *Myc* en la etiología de la enfermedad, quisimos confirmar mediante inmunohistoquímica la expresión de *Myc* en las células tumorales de los ratones transgénicos. Todos los tumores mostraron tinción para *Myc*, pero esta tinción estuvo ausente en las células del bazo de los ratones control *Cbp<sup>WT/F</sup> x EμBcl2* de la misma edad. Se muestran ejemplos ilustrativos, con tinción inespecífica en algunas células del bazo control (*Cbp<sup>WT/F</sup> x EμBcl2*), pero con una fuerte tinción positiva en los bazos tumorales de ratones con uno o ambos alelos de *Crebbp* suprimidos.

Estos datos, por lo tanto, apoyan un papel potencial de *Myc* en la regulación de los cambios en H3K18Ac que observamos en las células B con *Crebbp* inactivado.

### **1.5 Presencia de diferentes patrones mutacionales de *CREBBP* en el linfoma folicular en comparación con el linfoma difuso de células B grandes.**

Debido al gran contraste de nuestro modelo animal con el linfoma folicular humano, postulamos que las mutaciones de *CREBBP* podrían tener un papel diferente en el FL comparado con el DLBCL. Para obtener una idea sobre esto, llevamos a cabo la secuenciación dirigida de 126 muestras tumorales de FL humano y de 140 muestras tumorales de DLBCL humano y combinamos estos datos con los datos publicados previamente para 184 FL y 134 DLBCL<sup>132,133</sup>. Así, observamos que existía una proporción significativamente mayor de mutaciones del dominio KAT (dominio lisina acetiltransferasa) de *CREBBP* en los FL en comparación con los DLBCL (Figura 23A-B). Específicamente, las mutaciones de la arginina 1446 representaban el 26% de las mutaciones de *CREBBP* en el FL, pero sólo el 4% en el DLBCL. Por el contrario, las mutaciones de *CREBBP* en los tumores procedentes de DLBCL tendían a ocurrir con mayor frecuencia en regiones aguas arriba del dominio KAT, presentando un ratio

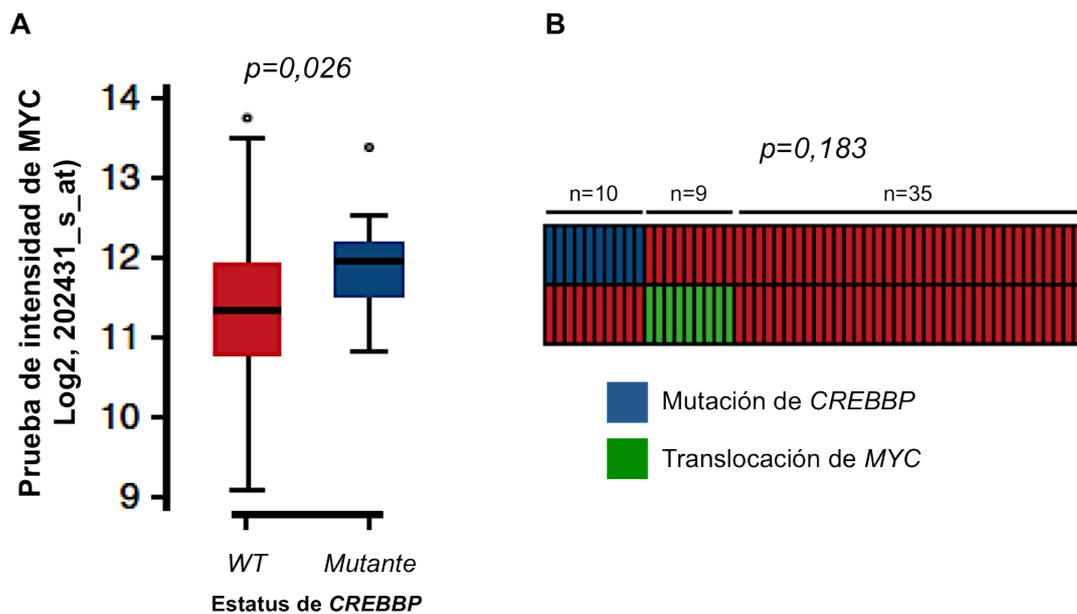
significativamente mayor de mutaciones *missense/nonsense* en comparación con el FL (**Figura 23C**).



**Figura 23: Distribución significativamente diferente de las mutaciones de *CREBBP* en el FL humano versus el DLBCL humano.** **A)** Se obtuvieron datos para la mutación de *CREBBP* de un total de 310 FL y 274 DLBCL humanos. Las mutaciones en *CREBBP* se identificaron en 151 FL y 51 DLBCL, y la representación relativa de la posición de estas mutaciones se expresa como una fracción de todas las mutaciones de *CREBBP* en esa enfermedad, en relación con la posición de la proteína *CREBBP* (isoforma 1). El dominio KAT está sombreado en verde y definido por los aminoácidos que van del 1342 al 1649. Picos más grandes, indicativos de una mayor fracción de mutaciones de *CREBBP*, se observan *aguas arriba* del dominio KAT en el DLBCL, en comparación con el FL. Además, un punto caliente dominante se puede ver en la arginina 1446 en el FL, que se reduce significativamente en el DLBCL. Por el contrario, otros puntos calientes en la tirosina 1482 y 1503 están presentes en frecuencias relativamente similares en FL y DLBCL. **B)** Los gráficos muestran que el 78% (118/151) de las mutaciones de *CREBBP* suceden dentro del dominio KAT en FL, en comparación con sólo el 43% (22/51) en DLBCL (test exacto de Fisher,  $p < 0,001$ ). **C)** Los gráficos muestran que sólo el 17% (25/151) de las mutaciones de *CREBBP* en FL presentan un cambio en el marco de lectura (*nonsense/frameshift*) creando un codón de parada prematura, mientras que el resto crea sustituciones de un solo aminoácido por inserciones/deleciones (*missense/in-frame*). Por el contrario, las mutaciones de cambio en el marco de lectura (*frameshift*) ocurren con más del doble de esta frecuencia, 39% (20/51), en los DLBCL (test exacto de Fisher,  $P = 0,0016$ ).

Tuvimos disponibles los análisis de expresión génica de 73 tumores procedentes de pacientes con DLBCL de los que se conocía su estatus para *CREBBP* (61 silvestre o *WT* y 12 mutantes). Aunque la translocación de *MYC* y el estado del número de copias de ADN no se conocían para estas muestras, sí pudimos determinar que los DLBCL con mutación de *CREBBP* mostraban un mayor expresión de *MYC* comparados con aquellos con *CREBBP* silvestre (*WT*) (**Figura 24A**). Por el contrario, tal y como publicaron nuestros colaboradores con anterioridad, no se observó una mayor expresión de *MYC* en asociación con las mutaciones de *CREBBP* en los FL<sup>143</sup>. Gracias a la colaboración con el grupo del Dr. *M. Green*, de la Universidad de Nebraska, se pudo analizar un grupo separado de 54 DLBCL procedentes del Centro

Médico de la Universidad de Nebraska (UNMC), en los que los datos sobre la translocación de *MYC* estaban disponibles (confirmados mediante hibridación fluorescente *in situ* o FISH), observándose que 9/54 de estos tumores eran positivos para la translocación *MYC*; además, 10/54 de los tumores presentaban mutaciones de *CREBBP*, y no hubo superposición entre estos elementos (**Figura 24B**). Sin embargo, esto no fue estadísticamente significativo, probablemente debido al tamaño limitado de la muestra (test exacto de Fisher,  $P = 0.183$ ).



**Figura 24: Estatus genético de *MYC* y *CREBBP* en muestras tumorales de pacientes con DLBCL.** **A)** Los análisis de expresión génica de tumores de pacientes con DLBCL, en los que previamente se había determinado el estatus mutacional de *CREBBP*, fueron usados para analizar la expresión de *MYC*. Se observó una expresión de *MYC* significativamente mayor en los tumores que poseían mutaciones de *CREBBP* en comparación con aquellos que no presentaban mutación de *CREBBP* (t test de *student*,  $P = 0.026$ ). Hay que tener en cuenta que en estas muestras el estatus de la translocación de *MYC* es desconocido y podría estar alterando la expresión de *MYC* en algunos casos. **B)** Un grupo de 54 nuevos casos de DLBCL fueron proporcionados por el UNMC. En ellos, el estatus de la translocación de *MYC* era conocido (9/54 presentaban dicha translocación). En ellos, también se detectó la mutación de *CREBBP* en 10/54. No se observó ninguna superposición entre la mutación *CREBBP* y la translocación de *MYC*, aunque esto no fue estadísticamente significativo.

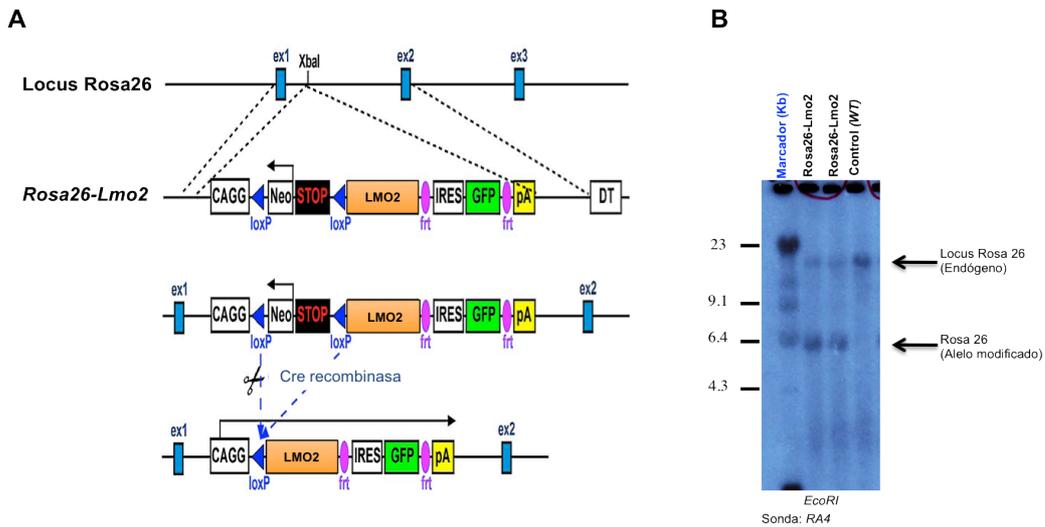
En conjunto, estos datos demuestran que existe un espectro significativamente diferente de las mutaciones de *CREBBP* en los DLBCL en comparación con los FL y proporcionan evidencias de que las mutaciones de *CREBBP* se asocian con un aumento de la expresión de *MYC* en los tumores de pacientes humanos con DLBCL.

### 2. La expresión transitoria de *Lmo2* en células madre hematopoyéticas (HSCs) causa T-ALL dependiente del timo.

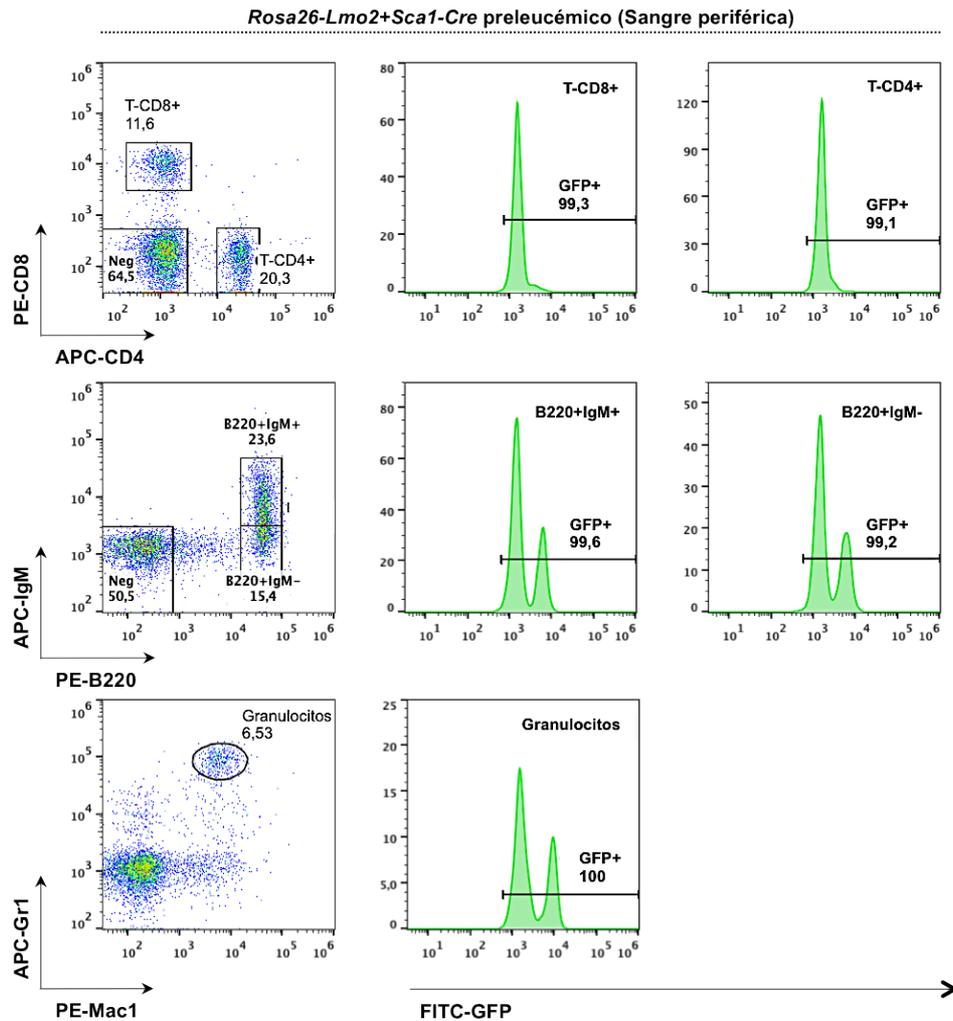
#### 2.1 Generación de una cepa de ratón en la que la expresión condicional de *Lmo2* está dirigida a HSCs.

La expresión de *LMO2* debida a la inserción y transactivación retroviral en células madre hematopoyéticas CD34+ de pacientes con X-SCID induce T-ALL, pero no otros tumores hematopoyéticos. Esta observación sugiere que *LMO2* tiene un efecto de reprogramación tumoral sobre las HSCs que lleva al inicio de la T-ALL, una idea apoyada por la habilidad de *LMO2* para inducir la regulación negativa de la diferenciación eritroide<sup>256</sup> y para generar células iPS<sup>95,259,292</sup>.

Por ello, nos planteamos generar un modelo condicional para *Lmo2* con el fin de estudiar su capacidad para reprogramar las HSCs. Con este objetivo, el ADNc de *Lmo2* fue introducido bajo el control del locus del promotor ubicuo Rosa26, unido al ADNc de *eGFP* a través de un sitio interno de entrada al ribosoma (*IRES-eGFP*). Para evitar la transcripción del ADNc de *Lmo2* y *eGFP* por el promotor Rosa26, una construcción que contenía el gen de resistencia a neomicina y un codón de parada flanqueado por sitios *loxP* se entroncaron entre el promotor Rosa26 y la secuencia *Lmo2-IRES-eGFP* (**Figura 25A**). Las células madre de los embriones generados (en las cuales se confirmó la presencia de *Lmo2*) fueron inyectadas dentro de blastocistos para generar quimeras. Los ratones resultantes fueron cruzados para generar una progenie con un alelo condicional de *Lmo2* transmisible a la línea germinal (designado como *Rosa26-Lmo2*) (**Figura 25B**). Así, en ausencia de la Cre recombinasa, ni *Lmo2* ni *eGFP* se expresan. Para estudiar la potencial contribución a la leucemogénesis de la expresión persistente de *Lmo2* en todas las células hematopoyéticas, iniciando dicha expresión en el compartimento de célula madre/progenitora hematopoyética, cruzamos los ratones *Rosa26-Lmo2* con ratones *Sca1-Cre*<sup>293</sup>. En primer lugar, se analizó mediante citometría de flujo la expresión de GFP de las células hematopoyéticas de sangre periférica de ratones preleucémicos, para confirmar que los ratones *Rosa26-Lmo2+Sca1-Cre* eliminaban el codón de parada e iniciaban la transcripción de *eGFP* y, por tanto, también de *Lmo2*. Así, observamos que todas las células de sangre periférica expresaban GFP, confirmando que existía expresión de GFP en todas las células hematopoyéticas (**Figura 26**).



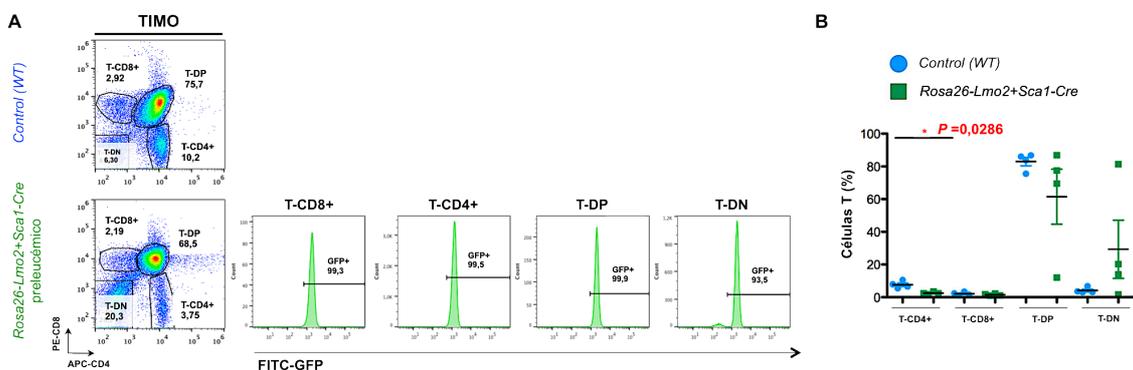
**Figura 25: Expresión condicional del oncogén *Lmo2* en HSCs de ratón.** A) Esquema que muestra el ADNc de *Lmo2* clonado dentro del locus *Rosa26* de ratón. Neo, casete de resistencia a la neomicina; poliA, secuencia de parada de poliadenilación; IRES, sitio interno de entrada al ribosoma; eGFP, proteína verde fluorescente. La recombinación mediada por la enzima Cre recombinasa elimina la secuencia de parada e inicia la expresión del oncogén *Lmo2* y de eGFP. B) Identificación de los ratones *Rosa26-Lmo2* mediante análisis de *Southern blot* del ADN de cola después de la digestión con *EcoRI*.



**Figura 26: Análisis de citometría de flujo representativo de la expresión de GFP en los diferentes compartimentos hematopoyéticos de la sangre periférica de ratones *Rosa26-Lmo2+Sca1-Cre* preleucémicos.** Como se muestra en la figura, todas las células hematopoyéticas de sangre periférica (células T, células B y granulocitos) de estos ratones expresan GFP, confirmando la eliminación del codón de parada por la Cre-recombinasa y la activación de la transcripción de *Lmo2* y eGFP.

## 2.2 La expresión de *Lmo2* en células madre hematopoyéticas de ratón causa exclusivamente T-ALL, similar a la T-ALL humana.

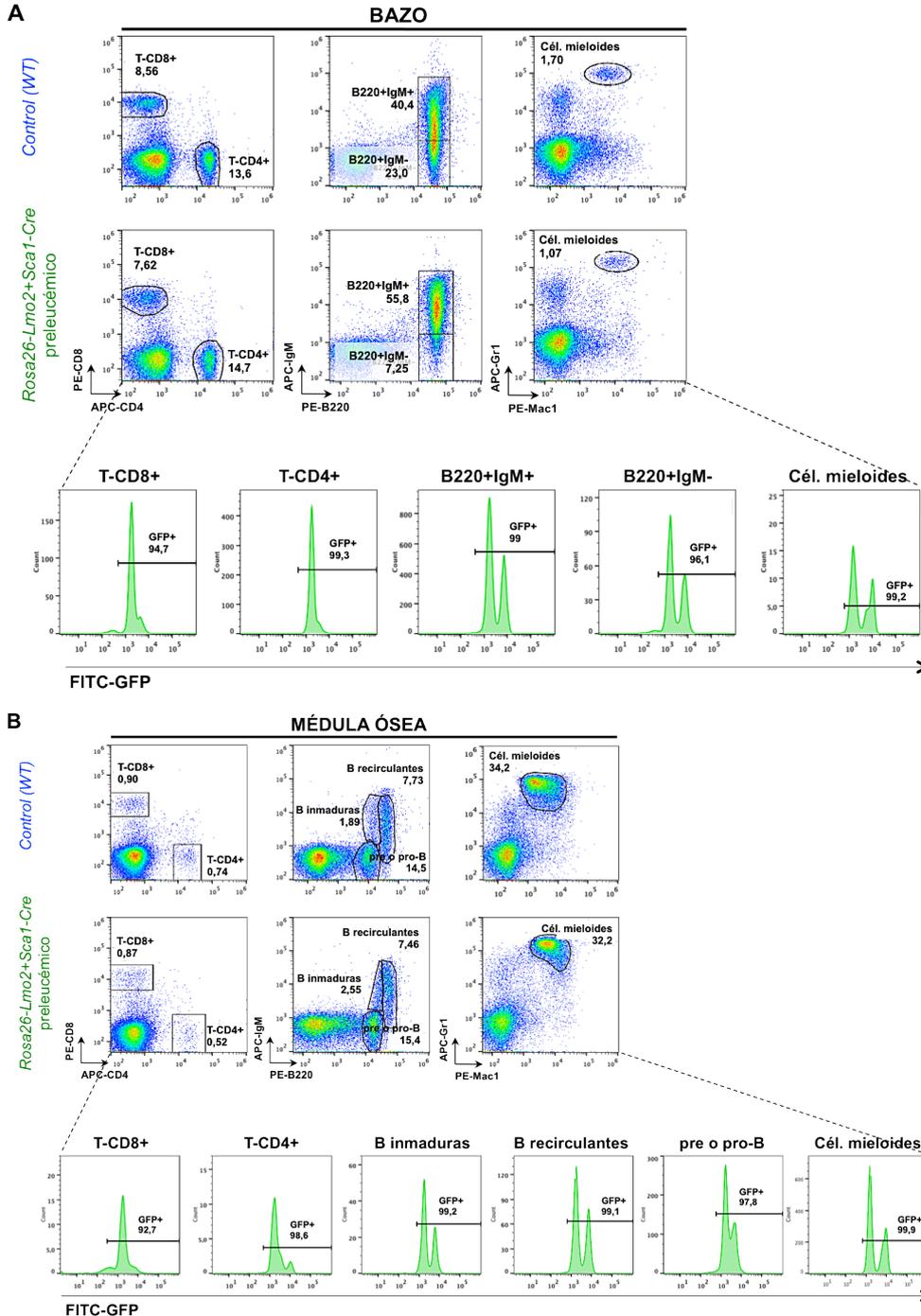
Para determinar si la expresión de *Lmo2* en HSCs predispone al desarrollo de T-ALL y demostrar así que esta expresión actuaría como el evento iniciador de las T-ALLs asociadas a terapia génica, monitorizamos una cohorte de ratones *Rosa26-Lmo2+Sca1-Cre* y en paralelo se estudió una cohorte de ratones control de tipo silvestre o *wild type* (*WT*). Para estudiar el impacto de *Lmo2* en la hematopoyesis a corto plazo, los diferentes compartimentos celulares hematopoyéticos fueron analizados en ratones *Rosa26-Lmo2+Sca1-Cre* jóvenes (de 4 meses de edad) mediante citometría de flujo. En primer lugar, quisimos determinar si la sobreexpresión de *Lmo2* en HSCs inducía algún defecto en el desarrollo de células T en los ratones en estudio. Para ello, se analizaron los timos de ratones *Rosa26-Lmo2+Sca1-Cre* y *WT* (de 4 meses de edad) mediante citometría de flujo, observándose que los ratones *Rosa26-Lmo2+Sca1-Cre* presentaban un porcentaje significativamente menor de células T-CD4+ en comparación con los ratones *WT* (Figura 27A-B; test de Mann Whitney,  $P = 0,0286$ ). Para el resto de subpoblaciones celulares T del timo no se observaron diferencias significativas entre cepas (test de Mann Whitney,  $P > 0.005$ ).



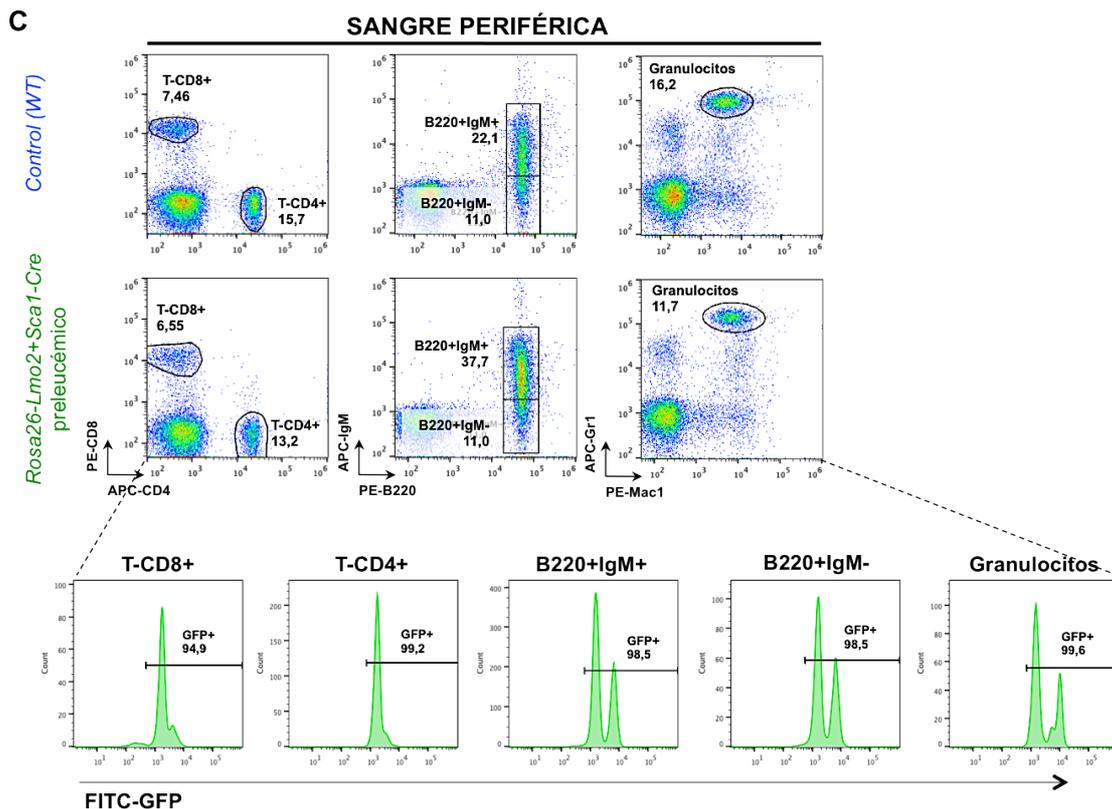
**Figura 27: Los ratones *Rosa26-Lmo2+Sca1-Cre* preleucémicos no presentan grandes alteraciones en el desarrollo T del timo como consecuencia de la expresión de *Lmo2* en HSCs. A)** Análisis de citometría de flujo del timo de un ratón preleucémico *Rosa26-Lmo2+Sca1-Cre* de 4 meses de edad. Se muestra un ejemplo representativo de la distribución de células T en el timo de estos ratones junto con la expresión del marcador GFP en los diferentes compartimentos celulares. El timo de un ratón control (*WT*) de edad similar se muestra como referencia. **B)** Porcentaje de células T analizadas mediante citometría de flujo en diferentes estadios del desarrollo (células T-CD4, células T-CD8, células T-DP-doble positivas y células T-DN-doble negativas) en los timos de ratones preleucémicos *Rosa26-Lmo2+Sca1-Cre* ( $n = 4$ ; 4 meses de edad) en comparación con ratones control *WT* ( $n = 4$ ; 4 meses de edad). Como se observa, los ratones *Rosa26-Lmo2+Sca1-Cre* preleucémicos muestran una reducción significativa (test de Mann Whitey,  $P = 0,0286$ ) en el porcentaje de células T-CD4+ del timo al compararlos con ratones control *WT* de la misma edad.

A continuación, se estudió si la expresión condicional de *Lmo2* en HSCs estaba induciendo algún defecto en el resto de compartimentos hematopoyéticos. A 4 meses de edad, los ratones condicionales para *Lmo2* no presentaban diferencias

significativas en los porcentajes de células B, T y mieloides de bazo, médula ósea y sangre periférica, al compararlos con los porcentajes que presentaban para esas poblaciones los ratones *control* (*WT*) de la misma edad (Figura 28A-C; test de Mann Whitney,  $P > 0,005$ ).



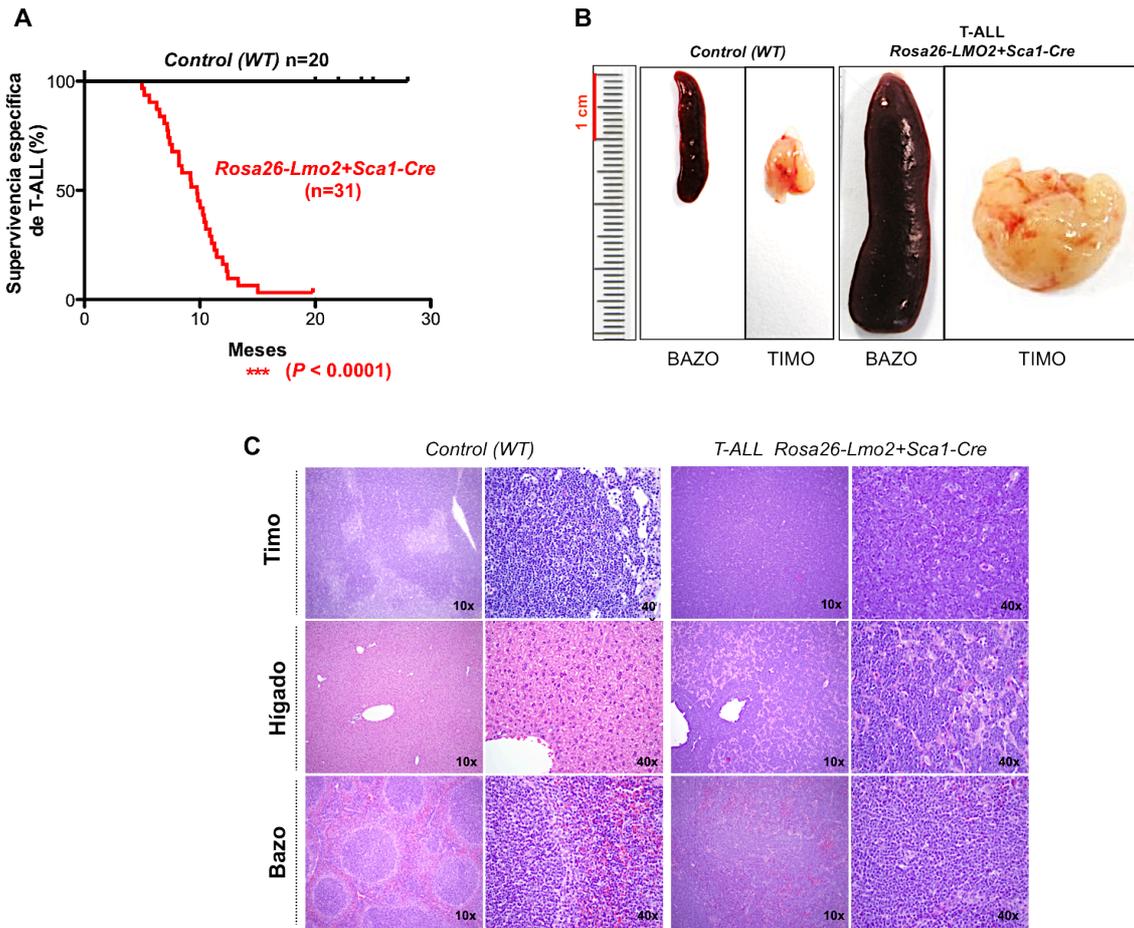
**Figura 28:** Los ratones *Rosa26-Lmo2+Sca1-Cre* preleucémicos muestran un desarrollo hematopoyético normal al compararlos con ratones *control* (*WT*) de la misma edad y expresan el marcador de seguimiento de *Lmo2* (GFP) en todas las células hematopoyéticas. **A-B** Análisis de citometría de flujo del bazo (**A**) y de la médula ósea (**B**) de un ratón preleucémico *Rosa26-Lmo2+Sca1-Cre* de 4 meses de edad. Se muestra un ejemplo representativo del desarrollo hematopoyético en el bazo y la médula ósea de estos ratones junto con la expresión del marcador GFP en los diferentes compartimentos celulares. El bazo y la médula ósea de un ratón *control* (*WT*) de edad similar se muestra como referencia. **Continúa en la siguiente página.**



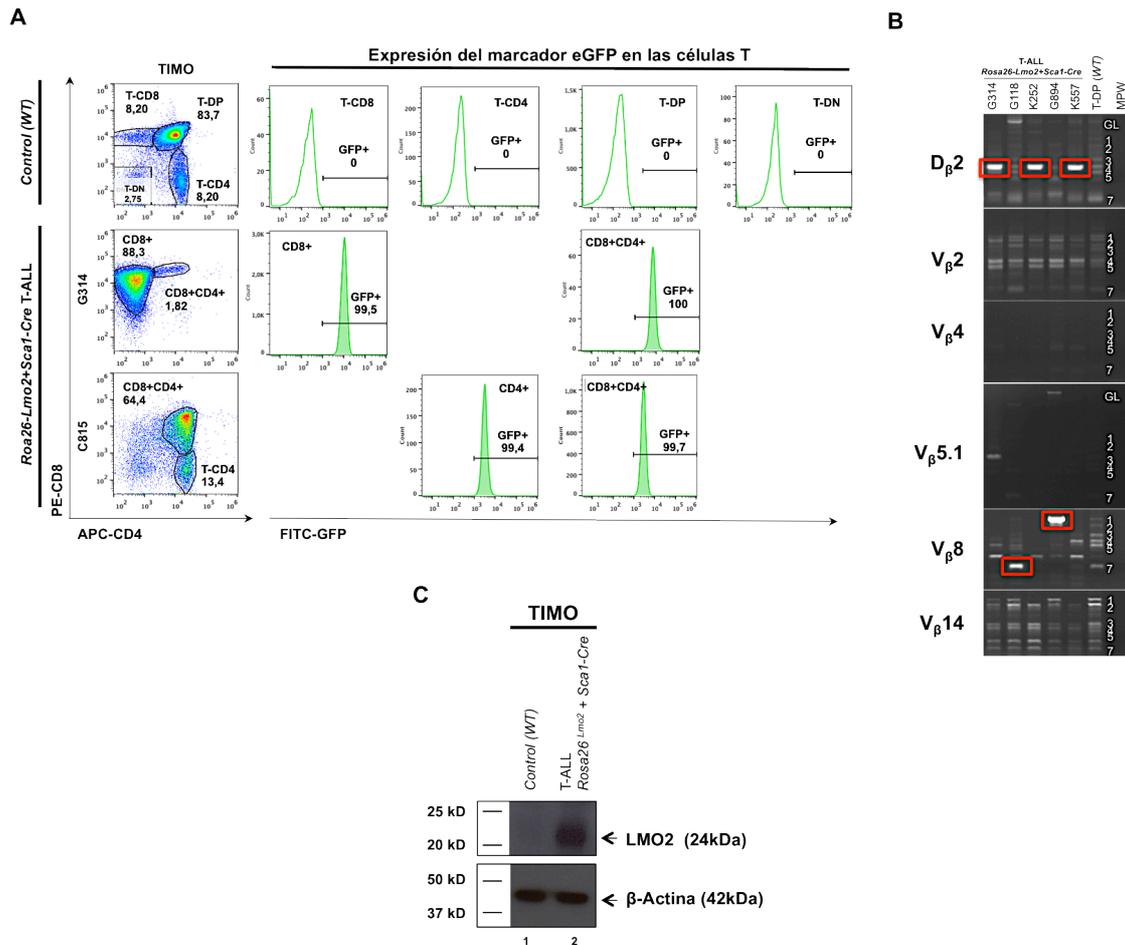
**Figura 28 (continuación): C)** Análisis de citometría de flujo de la sangre periférica de un ratón preleucémico *Rosa26-Lmo2+Sca1-Cre* de 4 meses de edad. Se muestra un ejemplo representativo del desarrollo hematopoyético en la sangre periférica de estos ratones junto con la expresión del marcador GFP en los diferentes compartimentos celulares. La sangre periférica de un ratón *control (WT)* de la misma edad se muestra como referencia.

En este escenario, los ratones fueron monitorizados a lo largo del tiempo, hasta que fueron sacrificados al detectar en ellos enfermedad. De esta manera, pudimos observar que los ratones *Rosa26-Lmo2+Sca1-Cre* (n=31) desarrollaban exclusivamente T-ALL (30/31), presentando una esperanza de vida significativamente más corta que los ratones *control (WT)*; n=20) (**Figura 29A**;  $P < 0,0001$ ; test de Mantel-Cox). Además, estos ratones con T-ALL solían mostrar en el momento del sacrificio timoma, esplenomegalia (**Figura 29B**) y pérdida de la arquitectura tisular del timo, hígado y bazo (**Figura 29C**).

Los análisis de citometría de flujo de las células leucémicas mostraron que éstas presentaban un inmunofenotipo  $CD8^+ / CD4^{+/-}$  (**Figura 30A**). Además, llevamos a cabo estudios de clonalidad, mediante PCR, que revelaron que estas leucemias presentaban reordenamientos clonales del receptor de células T (*TCR*) (**Figura 30B**). Por último, confirmamos que las células tumorales de estas leucemias expresaban LMO2 (**Figura 30C**).



**Figura 29: Los ratones *Rosa26-Lmo2+Sca1-Cre* desarrollan T-ALL agresivas.** **A)** Curva de supervivencia específica para T-ALL. Los ratones *Rosa26-Lmo2+Sca1-Cre* (línea roja, n=31) muestran una esperanza de vida significativamente más corta (test Mantel-Cox, *log-rang*;  $p < 0,0001$ ) comparada con la de los ratones *control (WT)* (línea negra; n=20) como consecuencia del desarrollo de T-ALL (30/31). **B)** Ejemplo representativo de esplenomegalia y timoma observado en los ratones *Rosa26-Lmo2+Sca1-Cre* con T-ALL. El bazo y el timo de un ratón *control (WT)* de edad similar se muestra como referencia. **C)** Tinción hematoxilina-eosina del timo, hígado y bazo de ratones leucémicos *Rosa26-Lmo2+Sca1-Cre* comparados con los de ratones *control (WT)* de la misma edad. Se puede observar como los tejidos de los ratones leucémicos *Rosa26-Lmo2+Sca1-Cre* pierden su arquitectura normal como consecuencia de la infiltración de células blásticas.



**Figura 30: Los ratones *Rosa26-Lmo2+Sca1-Cre* desarrollan T-ALL CD8<sup>+</sup> / CD4<sup>+</sup>, clonales y cuyas células tumorales expresan LMO2. A)** El análisis mediante citometría de flujo del inmunofenotipo de las T-ALL de los ratones *Rosa26-Lmo2+Sca1-Cre* nos permitió confirmar que estas leucemias se caracterizaban por ser CD8<sup>+</sup> / CD4<sup>+</sup>. Se muestran dos tipos representativos de ratones *Rosa26-Lmo2+Sca1-Cre* con T-ALL. Además, también se muestra la expresión de GFP de las células leucémicas. El timo de un ratón *control (WT)* se muestra como referencia. **B)** Clonalidad del receptor de células T (*TCR*) en ratones *Rosa26-Lmo2+Sca1-Cre*. Se muestra el análisis mediante PCR de los reordenamientos del gen *TCR* en timos infiltrados de ratones *Rosa26-Lmo2+Sca1-Cre* leucémicos. Las células T-DP aisladas del timo de un ratón *WT* sano fueron usadas como control de los reordenamientos policlonales del *TCR*. Los cuadros rojos marcan la clonalidad de los timos tumorales de los ratones *Rosa26-Lmo2+Sca1-Cre* leucémicos, observándose un incremento en la clonalidad del *TCR*. **C)** Análisis mediante *western blot* de la expresión de la proteína LMO2 en los timos tumorales de los ratones *Rosa26-Lmo2+Sca1-Cre* leucémicos (2). Se utilizó el timo de un ratón *WT* como control negativo (1). Como se puede observar, las células tumorales del timo de los ratones *Rosa26-Lmo2+Sca1-Cre* presentan una fuerte expresión de la proteína LMO2 que no se observa en el timo de los ratones *control (WT)*. Se muestra la expresión de la  $\beta$ -actina como control de carga. Los resultados son representativos de experimentos hechos por triplicado.

Para explorar la relevancia de nuestros resultados en las T-ALL humanas, llevamos a cabo la secuenciación del exoma completo (WES) de las células leucémicas de los ratones *Rosa26-Lmo2+Sca1-Cre* con T-ALL (n=9, **Tabla V**). Así, identificamos mutaciones recurrentes de variación de un solo nucleótido (SNVs) de *Notch1* (3/9), mutaciones de tipo indel también de *Notch1* (4/9), SNVs de *Kras* (3/9) y SNV de *Nras* (1/9) (**Tabla V**).

**Tabla V: Resumen del análisis de la secuenciación del exoma completo de las T-ALL de los ratones *Rosa26-Lmo2+Sca1-Cre* (n=9).**

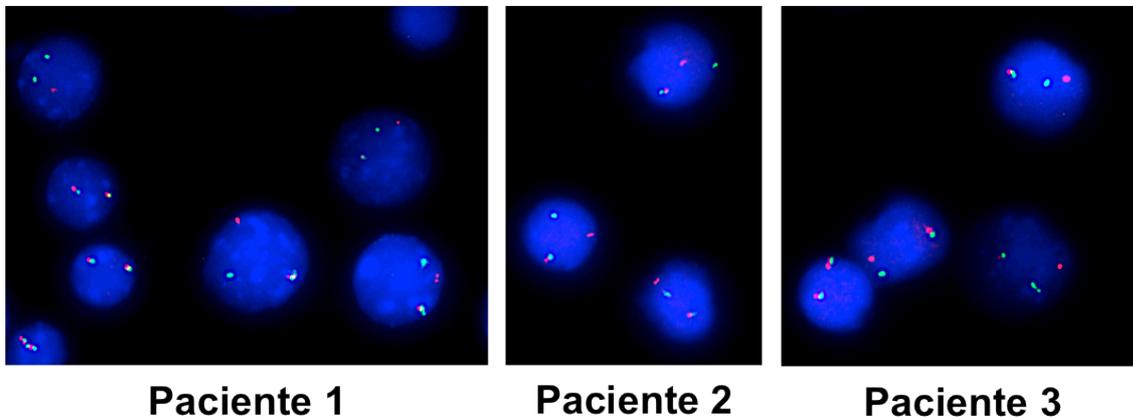
Análisis WES de las T-ALL de los ratones <i>Rosa26-Lmo2+Sca1-Cre</i>						
	Código del ratón	Edad de sacrificio (meses)	Notch1 SNV*	Notch1 Indel*	Kras SNV*	Nras SNV*
1	G314	5,17	-	p.2398	p.G12D	-
2	C815	9,23	-	p.2361	-	-
3	C812	10,37	-	p.2360	-	-
4	G894	10,03	p.R2361P	-	-	p.Q61K
5	K252	12,33	-	-	-	-
6	K557	7,4	p.L1585P	-	-	-
7	K563	10,26	-	p.2441	-	-
8	K565	10,53	-	-	p.K117N	-
9	K904	6,5	p.T2441P	-	p.Q61R	-

\*: mutaciones confirmadas mediante secuenciación por el método *Sanger*

Gracias a la colaboración con el grupo del Dr. A. *Borkhardt*, del Hospital de Dusseldorf, pudimos disponer de una cohorte de T-ALL humanas *LMO2+* (n=9) en las que la translocación t(11;14)(q13;q11) fue confirmada mediante análisis de hibridación fluorescente *in situ* (FISH) (**Figura 31**). La secuenciación mediante WES de estas T-ALL humanas *LMO2+* confirmó la importancia de las mutaciones encontradas en las T-ALL murinas. Así, encontramos 20 mutaciones somáticas, que incluían a 7 genes asociados a cáncer, confirmando mutaciones recurrentes de *NOTCH1* (SNVs 5/9, indel 1/9), *KRAS* (SNVs 1/9), *NRAS* (SNVs 2/9), *FBXW7* (SNVs 2/9) y *MTOR* (SNVs 2/9) (**Tabla VI**). Tanto las T-ALL humanas *LMO2+* como las de los ratones *Rosa26-Lmo2+Sca1-Cre* mostraron con mucha frecuencia mutaciones de *NOTCH1* y de *KRAS*,

afectando algunas de estas mutaciones al mismo aminoácido en ambas especies (Tabla VII).

## T-ALL humana t(11;14) (p13;q11)



**Figura 31: Análisis mediante FISH de pacientes con T-ALL y translocación t(11;14) positiva.** La co-localización de la señal rojo/verde define el alelo *LMO2* de tipo salvaje, mientras que las señales rojas y verdes por separado proporcionan una evidencia indirecta de la presencia de la translocación que genera el gen de fusión *TRD/LMO2*. Se muestran 3 ejemplos representativos de los 8 pacientes analizados.

**Tabla VI: Resumen del análisis de la secuenciación del exoma completo de las T-ALL humanas t(11;14)(p13;q11); n=9.**

Análisis WES de las T-ALL humanas t(11;14) (p13;q11)									
	Código de muestra	Edad del diagnóstico (años)	Tejido	NOTCH1 SNV*	NOTCH1 Indel*	KRAS SNV*	NRAS SNV*	FBXW7 Humano*	MTOR SNV*
1	60105	6.4	SP	-	-	-	-	-	-
2	100292	9.3	SP	p.L1585P	p.2514	-	-	-	p.V3006F
3	100300	13.3	MO	-	-	-	-	p.R289H	-
4	100386	9.8	MO	p.L1593P	-	-	p.G12S	p.R289H p.R46*	-
5	110289	18	SP	-	-	-	-	-	p.T2500F
6	110338	8.7	SP	p.L1596H	-	-	-	-	-
7	110297	2.3	MO	p.L1678P	-	-	p.G12D	-	-
8	120444	3	SP	-	-	-	-	-	-
9	120447	4.6	MO	p.P1377S	-	p.G12V	-	-	-

\*: mutaciones confirmadas mediante secuenciación por el método *Sanger*

Tabla VII: Tabla comparativa de los resultados de WES en T-ALL de ratones *Rosa26-Lmo2+Sca1-Cre* y en pacientes humanos con T-ALL y translocación t(11;14)(p13;q11).

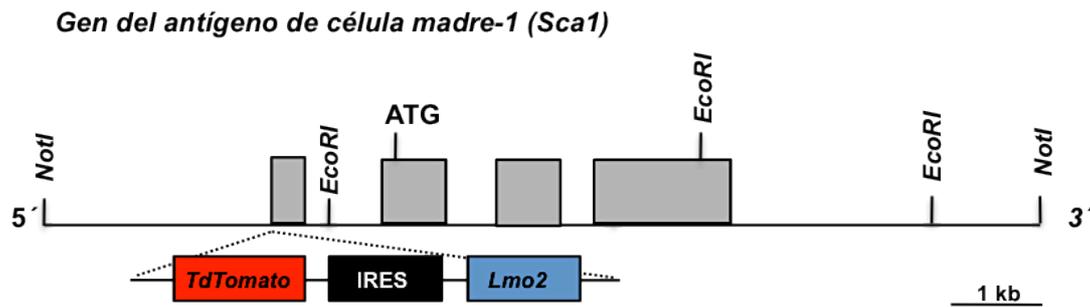
Genes	Mutaciones	<i>Rosa26-Lmo2+Sca1-Cre</i> (T-ALL)	T-ALL humanas t(11;14)(p13;q11)
<b>Notch1</b>	SNVs	<b>+(3/9)</b> c.863A>G; <b>p.L1585P</b> c.806T>G; p.T2441P c.045C>G; p.R2361P	<b>+(5/9)</b> c.768A>G; p.L1678P c.356A>T; p.L1596H c.365A>G; p.L1593P c.389A>G; <b>p.L1585P</b> c.219G>A; p.P1377S
	Indels	<b>+(4/9)</b> c.932c>CG; p.2398 c.043G>GC; p.2361 c.046G>GC; p.2360 C.804T>TGG; p.2441	<b>+(1/9)</b> c.648CAG>C; p.2514
<b>Kras</b>	SNVs	<b>+(3/9)</b> c.771C>T; <b>p.G12D</b> c.356T>C; p.Q61R c.194>G; p.K117N	<b>+(1/9)</b> c.284C>A; <b>p.G12V</b>
<b>Nras</b>	SNVs	<b>+(1/9)</b> c.270C>A; p.Q61K	<b>+(2/9)</b> c.747C>T; p.G12D c.748C>T; p.G12S

Se muestran en *negrita* aquellas mutaciones que afectan al mismo aminoácido en ambas especies.

En conjunto, estos resultados sugieren que la expresión temprana de *Lmo2* en el compartimento HSC/PCs es capaz de inducir una leucemia linfoblástica aguda de células T similar a la humana, desde un punto de vista inmunofenotípico, molecular y transcripcional.

### 2.3 *Lmo2* funciona como un oncogén “hit-and-run” en el desarrollo de la T-ALL.

Nuestros resultados previos sugieren que *Lmo2* es capaz de reprogramar tumoralmente a las HSCs hacia un fenotipo de T-ALL. Asimismo, se sabe que *Lmo2* está implicado en la reprogramación de células sanguíneas comprometidas de ratón, induciendo su transformación hacia HSCs<sup>259</sup>. Por ello, quisimos explorar si *Lmo2* podía funcionar como un oncogén “hit-and-run” dirigiendo la diferenciación de las HSCs hacia un fenotipo maligno de célula T, a través de un mecanismo de reprogramación tumoral. Este escenario fue modelado *in vivo*, mediante la generación de una línea de ratón transgénico en que la expresión de *Lmo2* se iniciaba en una HSC/PC, a través de la introducción del ADNc de *Lmo2* junto con el de un marcador de fluorescencia rojo de seguimiento (*Lmo2-TdTomato*) bajo el control del promotor *Sca1* que se expresa en HSCs (**Figura 32**). Así, se generó la cohorte de ratones transgénicos *Sca1-TdTomato-IRES-Lmo2*, a la que llamaremos *Sca1-Lmo2*.

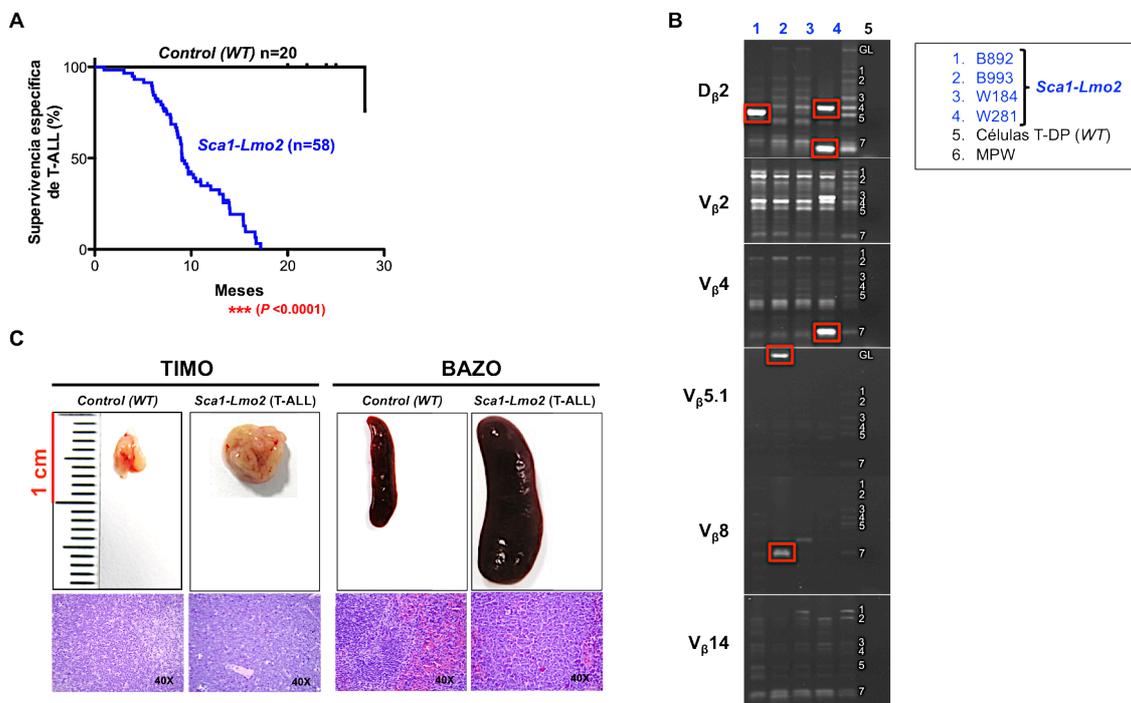


**Figura 32: Generación del ratón *Sca1-tdTOMATO-IRES-Lmo2*.** El ADN codificante del gen *Lmo2* murino se clonó en el primer exón del fragmento genómico que contiene el promotor y los exones del gen *Sca1* (Ly-6E) junto con el gen codificante de la proteína fluorescente roja TdTomato, unidos por una secuencia IRES (sitio interno de entrada al ribosoma).

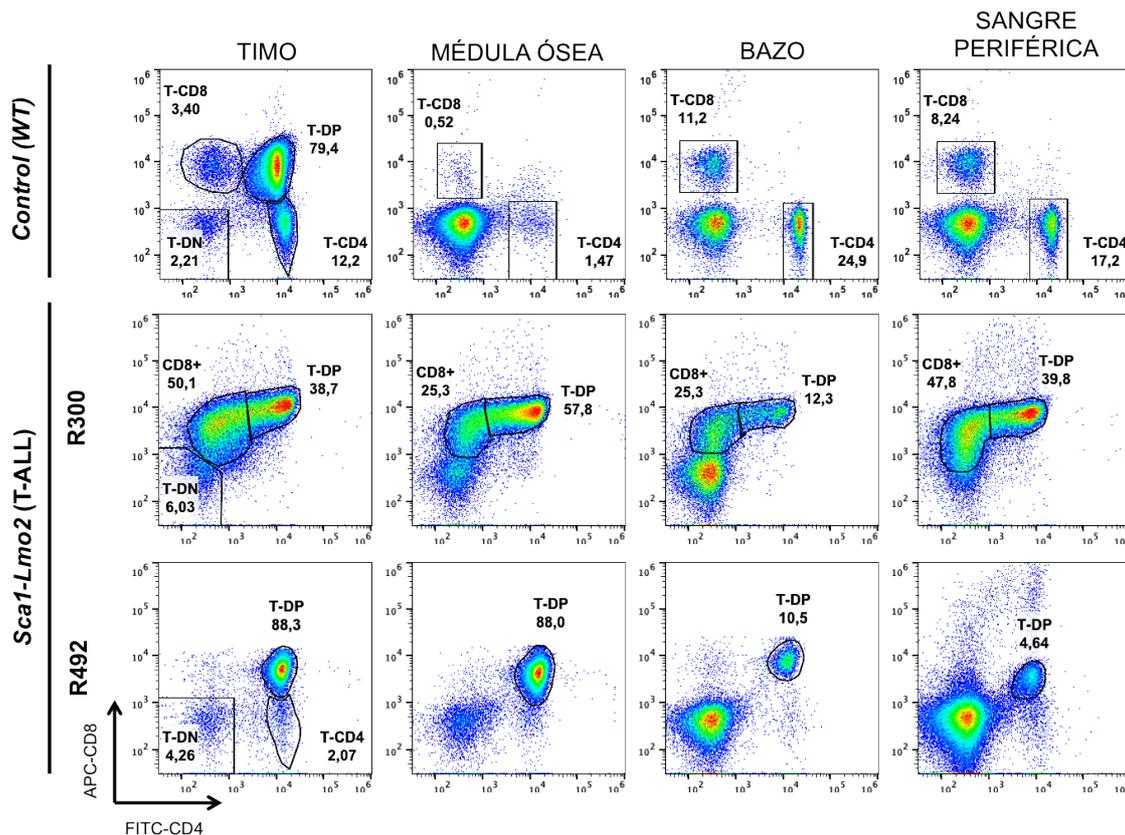
Estudios preliminares, presentados en la tesis doctoral de Isabel Romero Camarero<sup>294</sup>, titulada *Reprogramación tumoral en neoplasias linfoides*, mostraron que los ratones *Sca1-Lmo2* jóvenes (de 3-4 meses de edad) presentaban un desarrollo hematopoyético normal, con expresión del marcador de seguimiento TdTomato en todos los compartimentos celulares hematopoyéticos. Sin embargo, a esta edad, sí se observó un defecto en el desarrollo T en el timo de los ratones *Sca1-Lmo2*, presentando una disminución significativa del porcentaje de células T-DP (test de Mann Whitney;  $P < 0,0001$ ) que iba acompañada de un aumento significativo en el porcentaje de células T-DN y T-CD8+ (test de Mann Whitney;  $P = 0,0002$ ), al compararlos con los porcentajes de células T del timo de ratones *control* (*WT*) de la misma edad<sup>294</sup>. Consecuentemente, para probar si efectivamente como consecuencia de la sobreexpresión de *Lmo2*, los timocitos de los ratones *Sca1-Lmo2* preleucémicos presentaban estrés oncogénico, se procedió a analizar por PCR cuantitativa (qPCR) los niveles del supresor tumoral *Cdkn2a* ó *INK4a/ARF*, que codifica para la proteína ARF. Se analizaron los timos de 4 ratones *Sca1-Lmo2*, de 3 meses de edad, la médula ósea de un ratón que había mostrado previamente estrés oncogénico, como control positivo, y el timo de un ratón control (*WT*), como control negativo. En todos los ratones *Sca1-Lmo2* se encontró expresión de *Cdkn2a*, concluyendo, así, que los timos de dichos ratones estaban siendo expuestos a estrés oncogénico debido a la sobreexpresión de *Lmo2*. Por otro lado, se observó, mediante qPCR, que cada subpoblación celular T del timo de los ratones *Sca1-Lmo2* mostraba una expresión en mosaico de *Lmo2*, de manera que las células T *Lmo2*<sup>pos</sup> (positivas) y *Lmo2*<sup>neg</sup> (negativas) coexistían en el timo de estos ratones, sugiriendo un efecto de arrastre de la expresión de *Lmo2* desde el compartimento de las HSCs<sup>294</sup>.

Por lo tanto, el modelo *Sca1-Lmo2* nos permitió explorar si los progenitores de células T se seleccionaban clonalmente a través de un mecanismo de reprogramación de la diferenciación celular dirigido por la sobreexpresión de *Lmo2*.

Por ello, monitorizamos una cohorte de ratones *Sca1-Lmo2* (n=58) y una cohorte de ratones *control* (*WT*; n=20), los cuales fueron sacrificados cuando observamos enfermedad en ellos. Así, detectamos que los ratones *Sca1-Lmo2* presentaban una esperanza de vida significativamente menor (test de Mantel Cox, *log rank*;  $P < 0,0001$ ) que los ratones *control* (*WT*) (**Figura 33A**), debido al desarrollo de T-ALL clonales (n=45/58; **Figura 33B**) y altamente diseminadas (**Figura 33C**), caracterizadas por un inmunofenotipo  $CD8^+ / CD4^{+/-}$  (**Figura 34**) similar al observado en las T-ALL humanas<sup>208</sup> y en las T-ALL de los ratones *Rosa6-Lmo2+Sca1-Cre*.



**Figura 33: Los ratones *Sca1-Lmo2* presentan una esperanza de vida significativamente menor que los ratones *control* (*WT*) como consecuencia del desarrollo de T-ALL agresivas y clonales. A)** Curva de supervivencia específica para T-ALL. Los ratones *Sca1-Lmo2* (línea azul, n=58) muestran una esperanza de vida significativamente más corta (test Mantel Cox, *log-rang*  $p < 0,0001$ ) comparada con la de los ratones *control* (*WT*; línea negra; n=20) como consecuencia del desarrollo de T-ALL (45/58). **B)** Clonalidad del receptor de células T (*TCR*) en ratones *Sca1-Lmo2*. Se muestra el análisis mediante PCR de los reordenamientos del gen *TCR* en timos infiltrados de ratones *Sca1-Lmo2* leucémicos. Las células T-DP aisladas del timo de un ratón *WT* sano fueron usadas como control de los reordenamientos policlonales del *TCR*. Los cuadros rojos marcan la clonalidad de los timos tumorales de los ratones *Sca1-Lmo2* leucémicos, observándose un incremento en la clonalidad del *TCR* en los mismos. **C)** Ejemplo representativo de timoma y esplenomegalia observados en los ratones *Sca1-Lmo2* con T-ALL. El timo y el bazo de un ratón *control* (*WT*), de edad similar, se muestran como referencia. Por debajo de las imágenes macroscópicas de los timos y bazos se muestra la tinción hematoxilina-eosina correspondiente. Se observa como los tejidos de los ratones leucémicos *Rosa26-Lmo2+Sca1-Cre* pierden su arquitectura normal como consecuencia de la infiltración de células blásticas.



**Figura 34:** Las T-ALL de los ratones *Sca1-Lmo2* muestran un fenotipo CD8<sup>+</sup> / CD4<sup>+</sup>. Imágenes de citometría de flujo representativas de las T-ALL desarrolladas por los ratones *Sca1-Lmo2*. Como se puede observar, las células T leucémicas infiltran el resto de tejidos hematopoyéticos (médula ósea, bazo y sangre periférica). Se muestra el timo, médula ósea, bazo y sangre periférica de un ratón *control (WT)* de edad similar como referencia.

Las secciones histológicas de los timos tumorales de los ratones *Sca1-Lmo2* se caracterizaron por presentar una expresión positiva para CD3 (marcador específico de células T) y para TdT (marcador específico de células inmaduras/blásticas) (**Figura 35**). Sorprendentemente, gracias a la detección del marcador de seguimiento TdTomato mediante citometría de flujo, pudimos comprobar que sólo aquellos progenitores de células T en los cuales la expresión de *Lmo2* se había perdido, eran seleccionados para su transformación leucémica en los ratones *Sca1-Lmo2*, mientras que sólo una pequeña fracción de células residentes mantuvieron la expresión de *Lmo2* (**Figura 36**).

*Sca1-Lmo2* T-ALL (TIMO)

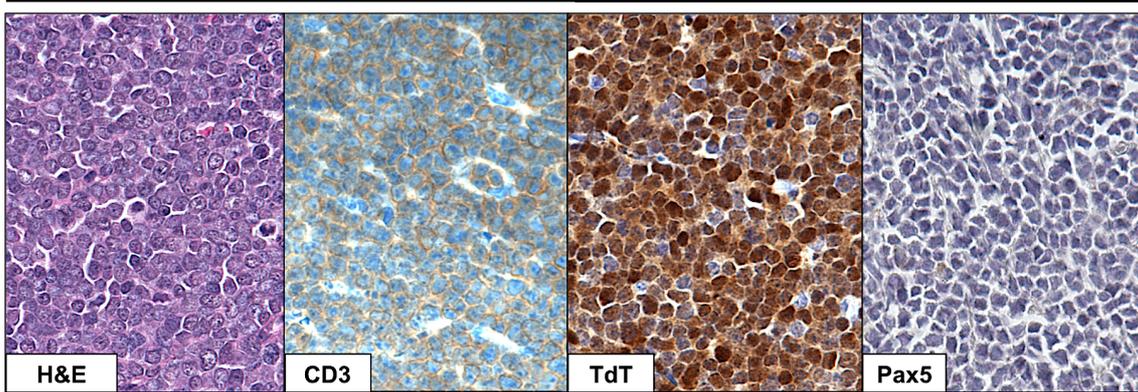


Figura 35: Secciones histológicas de inmunohistoquímica del timo de un ratón *Sca1-Lmo2* con T-ALL. Se muestra un ejemplo representativo del timo de un ratón *Sca1-Lmo2* leucémico teñido con hematoxilina-eosina y con los marcadores CD3 y TdT (correspondiente a una masa de células T blásticas) y Pax5 negativas (confirmando que no se trata de una neoplasia de células B). Los resultados son representativos de experimentos hechos por triplicado. Las imágenes tienen un tamaño de 300X aumentos.

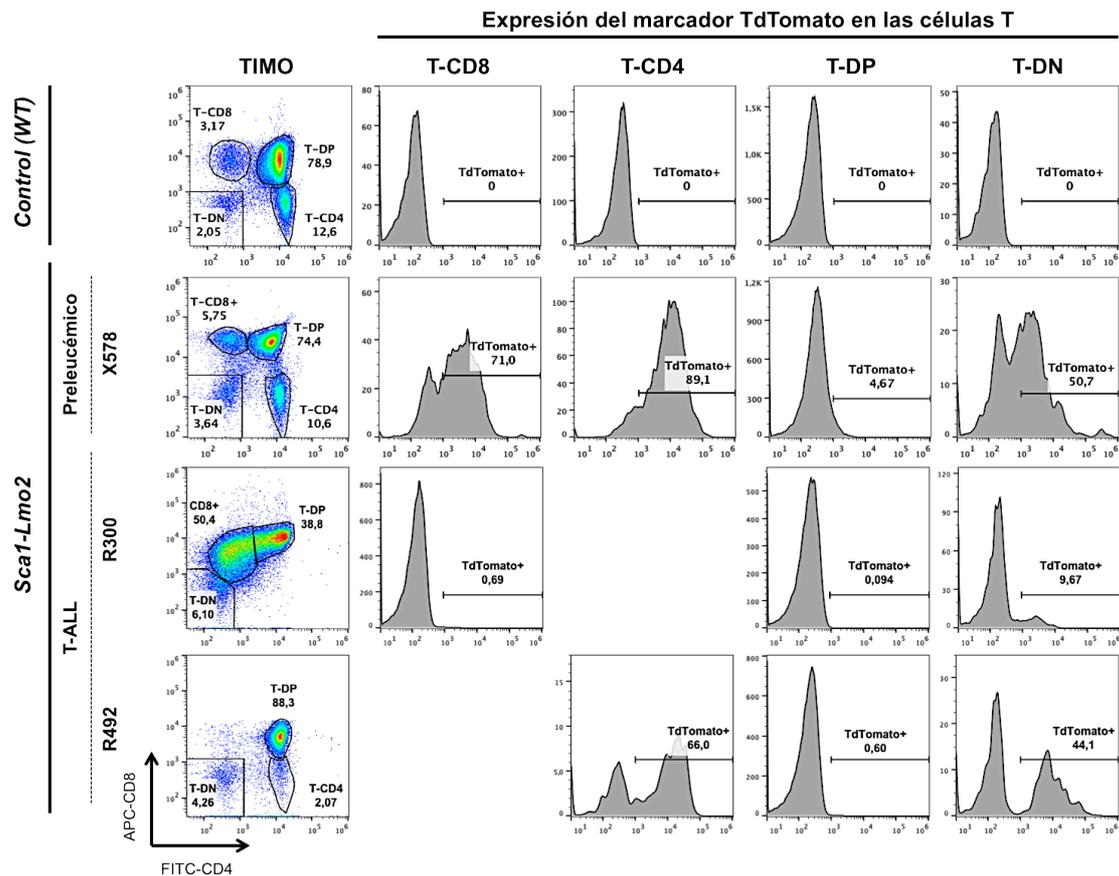


Figura 36: Únicamente los progenitores T que han perdido la expresión de *Lmo2* son seleccionados para su transformación tumoral en los ratones *Sca1-Lmo2*. Se muestran ejemplos de citometría de flujo de los timos de un ratón *Sca1-Lmo2* preleucémico y de dos ratones *Sca1-Lmo2* con T-ALL. Al lado se muestra la expresión del marcador TdTomato (marcador de seguimiento de la expresión de *Lmo2*), detectada mediante citometría de flujo, en los diferentes compartimentos celulares T del timo de estos ratones. Como se puede observar, las células T leucémicas de los ratones *Sca1-Lmo2* son TdTomato negativas en el 100% de los ratones *Sca1-Lmo2* analizados. Sin embargo, se mantiene una expresión en mosaico de *Lmo2* dentro de las poblaciones celulares T no tumorales. Se muestra el timo y la expresión de TdTomato en la diferentes poblaciones T de un ratón *control (WT)* y de un ratón *Sca1-Lmo2* preleucémico como referencia.

Todos estos datos indican que una expresión temprana del oncogén *Lmo2* en HSC/PCs es capaz de inducir T-ALL agresivas sin necesidad de mantener la expresión continua del oncogén en la masa tumoral.

## 2.4 Las células T leucémicas de los ratones *Sca1-Lmo2* muestran alteraciones genéticas y moleculares análogas a las que aparecen en las T-ALL humanas.

En las T-ALL humanas y murinas, las pérdidas y ganancias de material genético reflejan inestabilidad genómica<sup>280,281,295,296</sup>. Para determinar si las T-ALL desarrolladas por los ratones *Sca1-Lmo2* presentaban inestabilidad genómica, 12 tipos de ratones *Sca1-Lmo2* leucémicos fueron analizados mediante hibridación genómica comparada (CGH). Los resultados de este análisis mostraron que las leucemias analizadas presentaban, entre otras, pérdidas en el número de copias de *Cdkn2a/b* (2/12) y de *Bcl11b* (4/12), así como amplificaciones de *Myc* (8/12) (Tabla VIII). Por lo tanto, los progenitores T en los cuales la expresión de *Lmo2* se perdía, presentaban una inestabilidad genómica similar a la que aparece en las T-ALL humanas<sup>297,298</sup>.

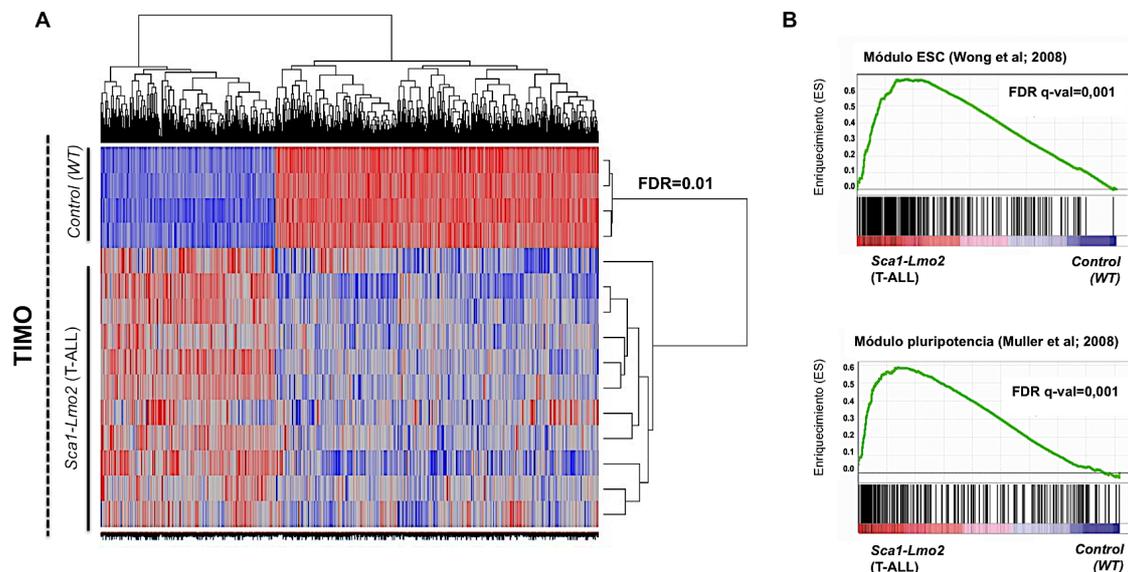
**Tabla VIII: Resumen de la inestabilidad genómica en las T-ALL de los ratones *Sca1-Lmo2* analizada mediante aCGH.**

Cromosoma (cr)	Citobanda	Tamaño (Mb)	Alteración	Frecuencia de las alteraciones	Genes
cr1	qD	0,161657	Delección	25%	<i>Alpi, Akp3, Ecel1, Chmd, Chmg, Eif4e2</i>
cr1	qD	0,097072	Duplicación	25%	<i>Ugt1a10, Ugt1a9, Ugt1a7c, Ugt1a6b, Ugt1a6a, Ugt1a5, Ugt1a2, DNajb3, Ugt1a1, Hijurp</i>
cr2	qE3	0,323916	Duplicación	25%	<i>Mpped2</i>
cr4	qC4 - qC5	1,852901	Delección homocigota	8,33%	<i>Mtap, Cdkn2a, Cdkn2b, Dmrt1, Zfp352, Ifna7, Ifna11, Ifna6, Ifna5, Ifna4, Ifna1, Ifne</i>
cr4	qC4 - qC5	3,008899	Delección	8,33%	<i>Mtap, Cdkn2a, Cdkn2b, Dmrt1, Zfp352, Elavl2</i>
cr4	qD1	2,114729	Duplicación	50%	<i>Skint4, Skint3, Skint9, Skint2, Skint10, Skint6, Skint5</i>
cr4	qD1	1,84909	Delección	25%	<i>Skint4, Skint3, Skint9, Skint2, Skint10, Skint6, Skint5</i>
cr6	qB1	0,470972	Delección homocigota	75%	<i>Trbv cluster, Try4, Try5, Try10, Prss3, EG436523, Prss1</i>
cr6	qB1	0,462656	Delección	33%	<i>Trbv cluster, Try4, Try5, Try10, Prss3, EG436523, Prss1</i>
cr11	qB4	0,117582	Delección homocigota	25%	<i>Nlrp1b, Nlrp1c</i>
cr11	qB4	0,132029	Duplicación	42%	<i>Nlrp1b, Nlrp1c</i>
cr11	qD	0,082762	Delección	16,67%	<i>Arhgap23, Scrin1</i>
cr11	qE2	0,170131	Delección	16,67%	<i>Cbx4, Tbc1d16</i>
cr12	qF1	0,065721	Delección	33%	<b><i>Bcl11b</i></b>
cr13	qC3	1,165686	Duplicación	25%	<i>Cox7c</i>
cr14	qC2	0,406833	Delección	50%	<i>Trav cluster, Trdv3, Trdv4, Trdd1, Trdd2, Trdj2, Traj cluster</i>
cr14	qC2	0,406833	Delección homocigota	66,67%	<i>Trav cluster, Trdv3, Trdv4, Trdd1, Trdd2, Trdj2, Traj cluster</i>
cr15	qA1 - qF3	100,318368	Trisomía	66,67%	1183 genes: <b><i>Myc</i></b>
cr17	qA3.3	0,433082	Duplicación	16,67%	<i>Btd9, Glo1, Dnahc8</i>
cr19	qC2	0,065376	Delección	16,67%	<i>Fgfbp3, Btaf1</i>

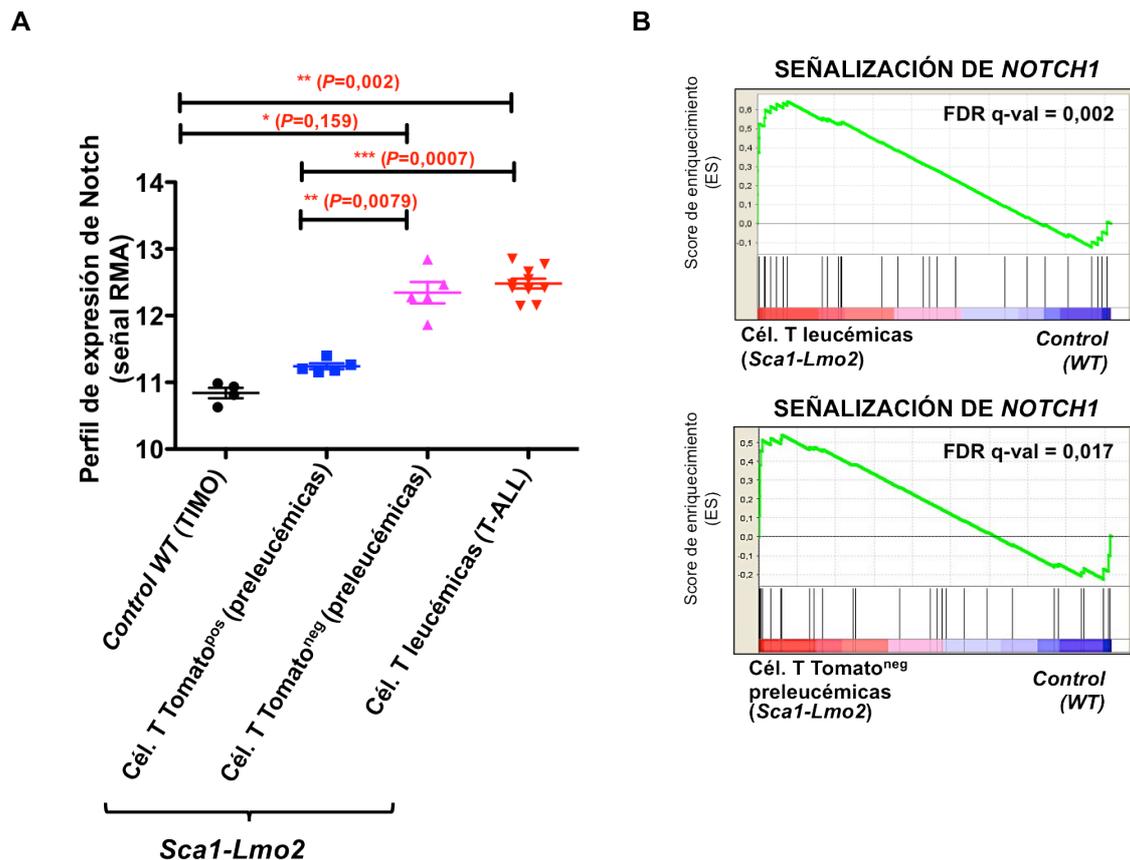
Marcadas en rojo aquellas alteraciones compartidas con las T-ALL humanas. cr: cromosoma; Mb: megabases.

Las T-ALL humanas se caracterizan por presentar alteraciones en importantes vías de señalización, incluyendo la activación de *NOTCH1*<sup>209</sup> y *MYC*<sup>299</sup>. Para explorar la

relevancia de nuestros resultados en la T-ALL humana, analizamos el perfil de expresión génica de las células leucémicas del timo de los ratones *Sca1-Lmo2*. Identificamos 6555 genes diferencialmente expresados (FDR=0,01) en los ratones *Sca1-Lmo2* leucémicos comparados con los ratones *WT* (**Figura 37A**). Entre los genes diferencialmente expresados en las células leucémicas de los ratones *Sca1-Lmo2*, encontramos la sobreexpresión de genes de la vía de señalización de *Notch1* y de dianas transcripcionales de *c-Myc*, así como una reducción de la expresión de *Fbxw7*, *Pten*, *Cyld* y *Cdkn1b*, los cuales también aparecen inactivados en las T-ALL humanas primarias<sup>210,300-306</sup>. Por lo tanto, las T-ALL *Lmo2*<sup>neg</sup> que surgen en los ratones *Sca1-Lmo2* presentan desregulación de las mismas rutas de señalización oncogénica que las T-ALL humanas. Además, gracias a los análisis de enriquecimiento, pudimos observar que las células tumorales *Lmo2*<sup>neg</sup> de los ratones *Sca1-Lmo2* leucémicos estaban altamente enriquecidas en genes asociados a células madre embrionarias (ESC) y a pluripotencia (**Figura 37B**), lo que sugiere que el programa transcripcional específico de célula madre permanece activado en ausencia de la expresión de *Lmo2*. Además, el análisis del perfil de expresión génica de células T preleucémicas de ratones *Sca1-Lmo2*, mostró que aquellas células preleucémicas que perdían la expresión de *Lmo2* presentaban una sobreexpresión de *Notch1* y una activación de su vía, comparable a la que se observaba en las células T leucémicas de estos ratones (**Figura 38A-B**).



**Figura 37: Las T-ALL de los ratones *Sca1-Lmo2* presentan un perfil de expresión génica similar al de las T-ALL humanas. A)** Diagrama de los 6555 genes expresados diferencialmente en las T-ALL de los ratones *Sca1-Lmo2* (n=10) en comparación con células T de ratones *control WT* (n=4), determinados usando un FDR=0,01. Cada columna representa un gen y cada fila representa una muestra. El nivel de expresión de cada gen se representa usando una escala de colores rojo-azul (los genes inducidos se representan en rojo y los genes reprimidos se representan en azul). **B)** Análisis de enriquecimiento (GSEA) de las T-ALL de los ratones *Sca1-Lmo2* comparados con células T de ratones *control WT*. Los GSEA mostraron que las T-ALL de los ratones *Sca1-Lmo2* estaban significativamente enriquecidas en genes asociados a células madre embrionarias (ESC; GSEA FDR q-valor=0,001) y a pluripotencia (GSEA FDR q-valor=0,001).



**Figura 38:** Las células T  $Lmo2^{neg}$  preleucémicas de los ratones *Sca1-Lmo2* presentan una sobreexpresión de *Notch1* comparable a la de las T-ALL desarrolladas por estos ratones. **A)** Perfil de expresión de *Notch1* en células T de ratones control WT (n=4), en células preleucémicas T  $Tomato^{pos}$  (n=5), en células preleucémicas T  $Tomato^{neg}$  (n=5) y en células T leucémicas (n=10) de ratones *Sca1-Lmo2*. Las células T  $Tomato^{neg}$  de los ratones *Sca1-Lmo2* preleucémicos muestran un perfil de expresión de *Notch1* similar al de las T-ALL de los ratones *Sca1-Lmo2* (t test de student). **B)** Análisis de enriquecimiento (GSEA) para la ruta de señalización de *Notch1* en células T leucémicas (GSEA FDR q-valor=0,002) y células T preleucémicas  $Tomato^{neg}$  (GSEA FDR q-valor=0,017) de ratones *Sca1-Lmo2* comparados con células T de ratones control WT. Las T-ALL y las células T preleucémicas  $Tomato^{neg}$  de los ratones *Sca1-Lmo2* están significativamente enriquecidas (FDR q-valor=0,002 y 0,017, respectivamente) en genes de la ruta de *Notch1*.

Por lo tanto, *Lmo2* es capaz de iniciar un mecanismo de reprogramación tumoral en las HSC/PCs, pero el fenotipo específico de T-ALL parece ser seleccionado independientemente de la expresión de *Lmo2*.

### 2.5 Las T-ALL $Lmo2^{neg}$ de los ratones *Sca1-Lmo2* comparten eventos secundarios críticos con las T-ALL humanas y las T-ALL de los ratones *Rosa26-Lmo2+Sca1-Cre*.

Con el objetivo de explorar la relevancia de nuestros resultados en la T-ALL humana asociada al tratamiento con terapia génica, llevamos a cabo WES de 10 T-ALL procedentes de los ratones *Sca1-Lmo2*. Así, detectamos 29 mutaciones somáticas,

que incluían mutaciones en 8 genes asociados a cáncer. Entre estas mutaciones, encontramos mutaciones recurrentes de *Notch1* (SNVs 5/10, Indels 5/10) y de *Kras* (SNVs 1/10) (**Tabla IX**), coincidiendo con las mutaciones que encontramos en las T-ALL humanas LMO2+ y en las T-ALL de los ratones *Rosa26-Lmo2+Sca1-Cre* (**Tabla X**). Sin embargo, no identificamos mutaciones en genes diana de *Lmo2* y/o en vías de señalización que pudieran reemplazar la función de *Lmo2* en las T-ALL *Lmo2*<sup>neg</sup>.

**Tabla IX: Resumen del análisis de la secuenciación del exoma completo de las T-ALL de los ratones *Sca1-Lmo2* (n=10).**

Análisis WES de las T-ALL de los ratones <i>Sca1-Lmo2</i>						
	Código del ratón	Edad de sacrificio (meses)	Notch1 SNV*	Notch1 Indel*	Kras SNV*	Nras SNV*
1	B892	10.47	-	p.2361	p.G12D	-
2	B993	9.7	-	p.2418	-	-
3	0899	6.4	p.L1668P	-	-	-
4	W184	7.83	p.R2361P	p.2362	-	-
5	W281	5.97	-	-	-	-
6	W283	6.27	-	p.2398	-	-
7	W759	9.7	-	-	-	-
8	C112	12.53	p.L1668P p.R2361P p.R1608P	p.2398	-	-
9	C120	12.67	p.R2361Q	-	-	-
10	C762	9.63	p.L1585P	-	-	-

\*: mutaciones confirmadas mediante secuenciación por el método *Sanger*

Tabla X: Tabla comparativa de los resultados de WES de las T-ALL de ratones *Rosa26-Lmo2+Sca1-Cre*, *Sca1-Lmo2* y de pacientes humanos con t(11;14)(p13;q11).

Genes	Mutaciones	<i>Rosa26-Lmo2+Sca1-Cre</i> (T-ALL)	<i>Sca1-Lmo2</i> (T-ALL)	T-ALL humanas t(11;14)(p13;q11)
<b>Notch1</b>	SNVs	<b>+(3/9)</b> c.863A>G; <b>p.L1585P</b> c.806T>G; p.T2441P c.045C>G; p.R2361P	<b>+(5/10)</b> c.863A>G; <b>p.L1585P</b> c.601A>G; p.L1668P c.045C>G; p.R2361P c.045C>G; p.R2361Q c.794C>G; p.R1608P	<b>+(5/9)</b> c.768A>G; p.L1678P c.356A>T; p.L1596H c.365A>G; p.L1593P c.389A>G; <b>p.L1585P</b> c.219G>A; p.P1377S
	Indels	<b>+(4/9)</b> c.932C>CG; p.2398 c.043G>GC; p.2361 c.046G>GC; p.2360 c.804T>TGG; p.2441	<b>+(5/10)</b> c.874A>ACC; p.2418 c.042A>AGAGG; p.2362 c.044C>CG; p.2361 c.932>CG; p.2398 c.934C>CCG; p.2398	<b>+(1/9)</b> c.648CAG>C; p.2514
<b>Kras</b>	SNVs	<b>+(3/9)</b> c.771C>T; <b>p.G12D</b> c.356T>C; p.Q61R c.194>G; p.K117N	<b>+(1/10)</b> c.771C>T; <b>p.G12V</b>	<b>+(1/9)</b> c.284C>A; <b>p.G12V</b>
<b>Nras</b>	SNVs	<b>+(1/9)</b> c.270C>A; p.Q61K	X	<b>+(2/9)</b> c.747C>T; p.G12D c.748C>T; p.G12S

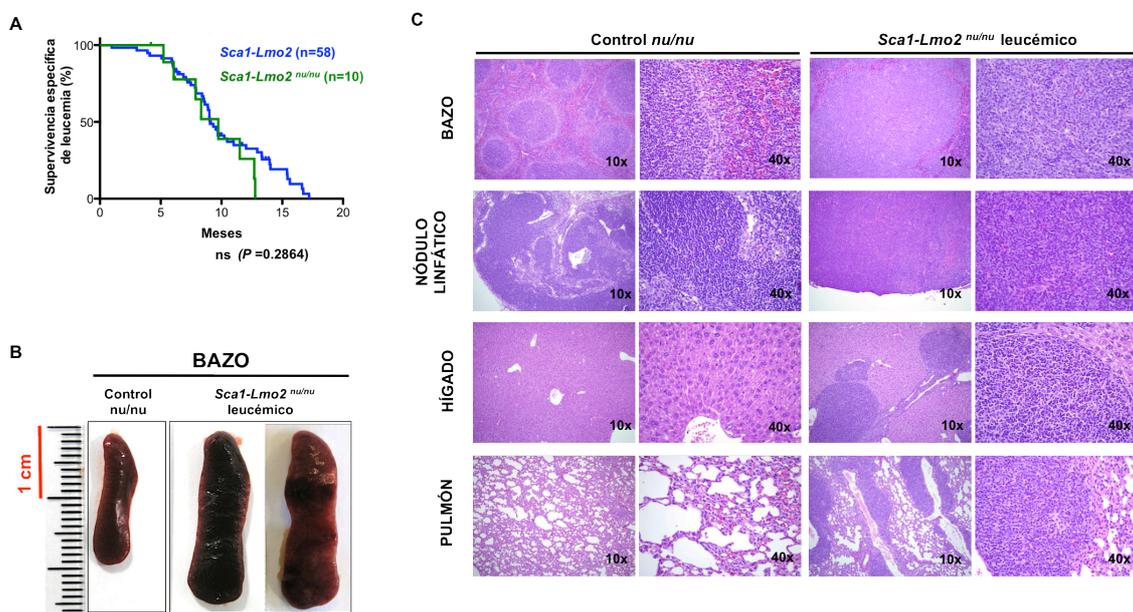
Se muestran en negrita aquellas mutaciones que afectan al mismo aminoácido en ambas especies.

Por lo tanto, nuestros datos sugieren que la expresión de *Lmo2* en HSCs sólo puede inducir T-ALL, mientras que la presión selectiva durante la maduración de las células T en el timo es necesaria para adquirir alteraciones genéticas secundarias, específicas de la T-ALL, que den lugar a la transformación final.

## 2.6 La adquisición de las alteraciones genéticas secundarias tiene lugar en el timo durante el desarrollo de las T-ALL.

Con el objetivo de determinar en qué ambiente (médula ósea o timo) estaba teniendo lugar el segundo evento oncogénico, cruzamos los ratones *Sca1-Lmo2* con ratones desnudos (*nu/nu*), carentes de timo. Así, generamos una cohorte de ratones *Sca1-Lmo2<sup>nu/nu</sup>* (n=10) que fue monitorizada junto con una cohorte de ratones *control nu/nu* (n=10), hasta que se detectó en ellos enfermedad. Los ratones *Sca1-Lmo2<sup>nu/nu</sup>* tenían una esperanza de vida similar a la de los ratones *Sca1-Lmo2*, como resultado del desarrollo de leucemias (**Figura 39A**). Los ratones *Sca1-Lmo2<sup>nu/nu</sup>* leucémicos (n=8/10) mostraron esplenomegalia y una forma de leucemia altamente diseminada que infiltraba tanto los tejidos hematopoyéticos como los tejidos no hematopoyéticos

(Figura 39B-C). Desde un punto de vista inmunofenotípico, los ratones *Sca1-Lmo2<sup>nu/nu</sup>* desarrollaron leucemias de células madre/progenitoras *Sca1+Lin+* (Figura 40). A diferencia de los ratones *Sca1-Lmo2*, las células leucémicas de los ratones *Sca1-Lmo2<sup>nu/nu</sup>* sí expresaban *Lmo2* (Figura 40). Se llevaron a cabo estudios de expresión génica diferencial de las leucemias de los ratones *Sca1-Lmo2<sup>nu/nu</sup>*, observando que presentaban 5801 genes diferencialmente expresados al compararlas con las T-ALL de los ratones *Sca1-Lmo2* (Figura 41A). Además realizamos análisis de enriquecimiento, que revelaron que estas leucemias estaban enriquecidas en genes asociados a células madre embrionarias (ESC) y pluripotencia (Figura 41B), poniendo en evidencia un efecto de reprogramación de *Lmo2* en las HSC/PCs.



**Figura 39: Los ratones *Sca1-Lmo2<sup>nu/nu</sup>* desarrollan leucemias agresivas. A)** Curva de supervivencia específica para leucemia. Los ratones *Sca1-Lmo2<sup>nu/nu</sup>* (línea verde, n=10) muestran una esperanza de vida similar a la de los ratones *Sca1-Lmo2* (línea azul, n=58; test Mantel-Cox, *log-rang*;  $P = 0.2864$ ) como consecuencia del desarrollo de leucemias. **B)** Ejemplos representativos de esplenomegalia observada en los ratones *Sca1-Lmo2<sup>nu/nu</sup>* con leucemia. El bazo de un ratón *control nu/nu* de edad similar se muestra como referencia. **C)** Tinción hematoxilina-eosina del bazo, nódulo linfático, hígado y pulmón de ratones leucémicos *Sca1-Lmo2<sup>nu/nu</sup>* comparada con la de ratones *control nu/nu* de edad similar. Se puede observar como los tejidos de los ratones leucémicos *Sca1-Lmo2<sup>nu/nu</sup>* pierden su arquitectura normal como consecuencia de la infiltración de células blásticas.

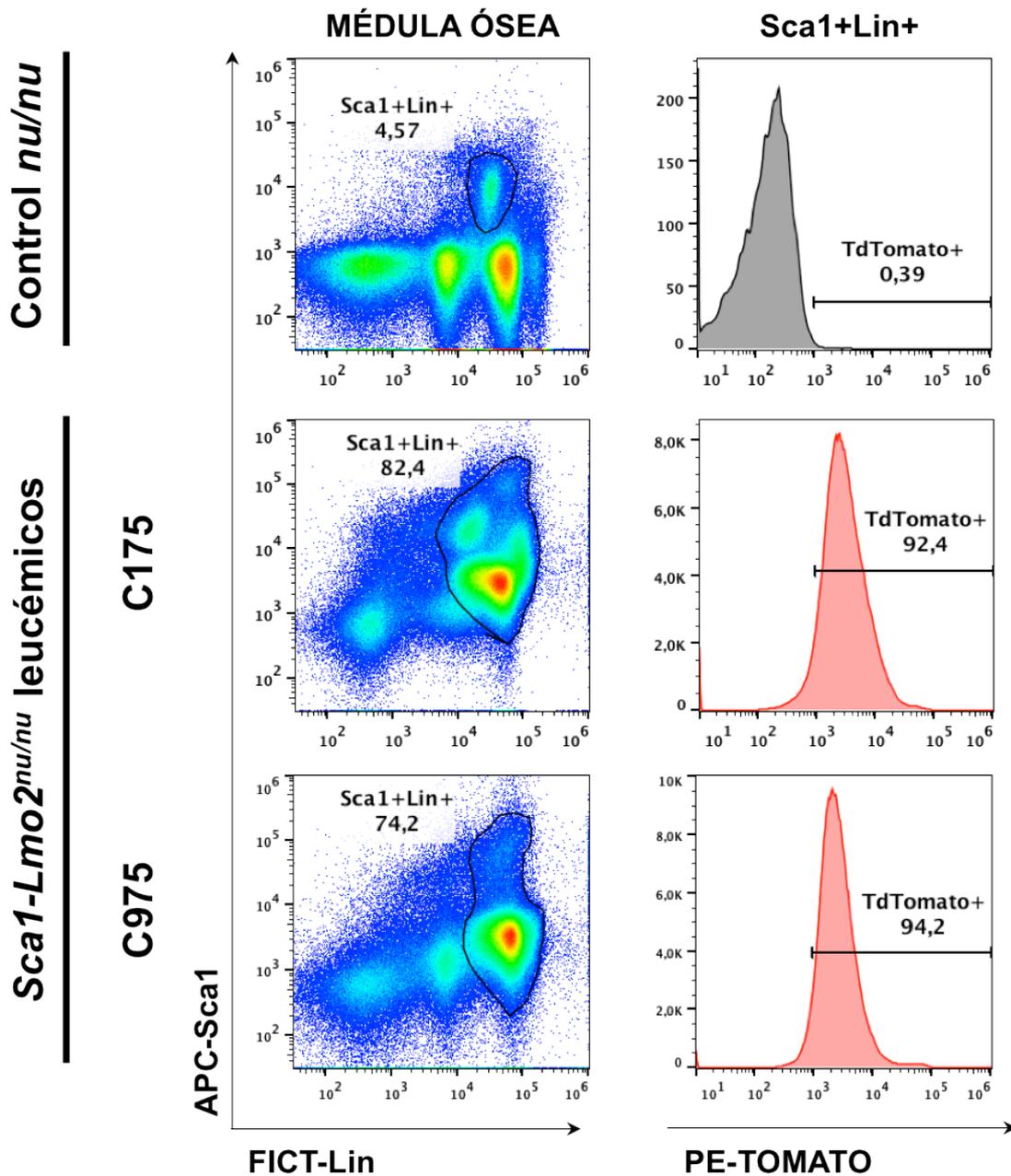
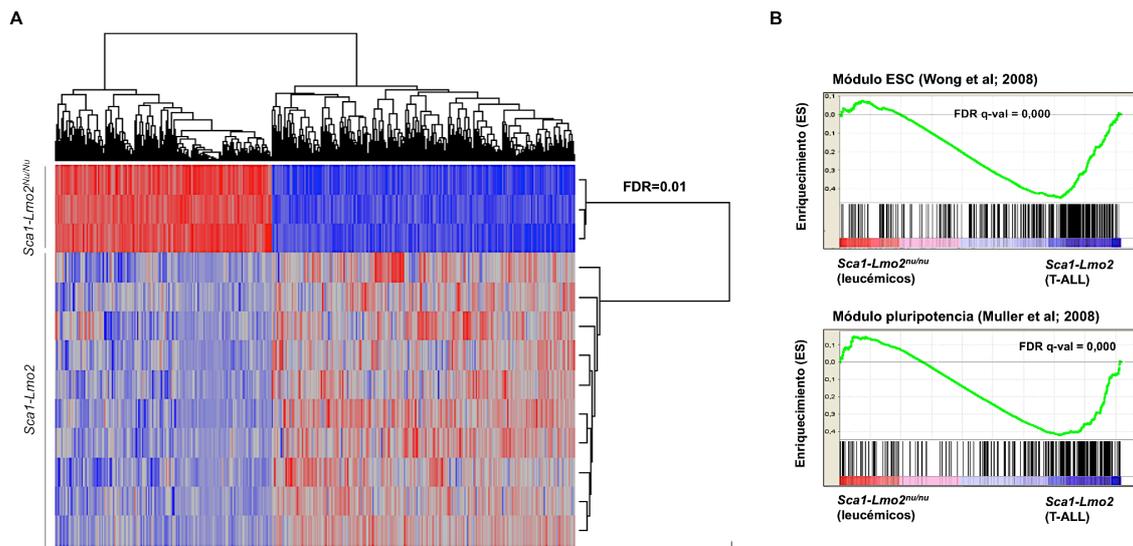


Figura 40: Los ratones *Sca1-Lmo2<sup>nu/nu</sup>* desarrollan leucemias con un inmunofenotipo inmaduro y expresión de *Lmo2*. Ejemplos representativos del análisis de citometría de flujo de la médula ósea de los ratones *Sca1-Lmo2<sup>nu/nu</sup>* leucémicos. Como se observa, estos ratones desarrollan masas de células Sca1+ Lin+. Además, confirmamos que estas leucemias expresaban *Lmo2* gracias a la detección en las células leucémicas del marcador de seguimiento TdTomato. La médula ósea de un ratón *control nu/nu*, de edad similar, se muestra como referencia. Lin=cóctel de linaje (CD4, CD8, CD3ε, CD19, B220, Mac1 y Gr1).



**Figura 41: Las leucemias de los ratones *Sca1-Lmo2<sup>nu/nu</sup>* muestran una firma genética enriquecida en genes asociados a células madre embrionarias y pluripotencia. A)** Diagrama de los 5801 genes expresados diferencialmente en la médula ósea de ratones *Sca1-Lmo2<sup>nu/nu</sup>* leucémicos (n=3) comparados con las T-ALL de los ratones *Sca1-Lmo2* (n=10; FDR=0,01). Cada columna representa un gen y cada fila representa una muestra. El nivel de expresión de cada gen se representa usando una escala de colores rojo-azul (los genes inducidos se representan en rojo y los genes reprimidos se representan en azul). **B)** Análisis de enriquecimiento (GSEA) de las leucemias de los ratones *Sca1-Lmo2<sup>nu/nu</sup>* comparadas con las T-ALL de los ratones *Sca1-Lmo2*. Los GSEA muestran que las leucemias de los ratones *Sca1-Lmo2<sup>nu/nu</sup>* están significativamente enriquecidas en genes asociados a células madre embrionarias (ESC; GSEA FDR q-valor=0,001) y a pluripotencia (GSEA FDR q-valor=0,001).

Estos resultados sugerían que el timo proporcionaba el ambiente necesario para la selección clonal de las HSC/PCs reprogramadas en ausencia de expresión de *Lmo2*. Por ello, llevamos a cabo WES de las leucemias de los ratones *Sca1-Lmo2<sup>nu/nu</sup>* (n=5), detectando 5 mutaciones somáticas recurrentes, de las cuales 3 afectaban a genes asociados a cáncer (*Spen*, *Tfdp1* y *Abi1*). Sin embargo, a diferencia de las T-ALL de los ratones *Rosa26-Lmo2+Sca1-Cre*, *Sca1-Lmo2* y de las T-ALL humanas LMO2+, no observamos mutaciones recurrentes de *Notch1* o *Ras* en las leucemias de los ratones *Sca1-Lmo2<sup>nu/nu</sup>* (Tabla XI), confirmando que el segundo evento oncogénico de las T-ALL estaba teniendo lugar en el timo.

Tabla XI: Resumen del análisis de la secuenciación del exoma completo de las leucemias de los ratones *Sca1-Lmo2<sup>nu/nu</sup>* (n=5)

Análisis WES de las leucemias de los ratones <i>Sca-Lmo2<sup>nu/nu</sup></i>							
	Código del ratón	Edad de sacrificio (meses)	Notch1 SNV*	Notch1 Indel*	Tfdp1 SNV*	Abl1 SNV*	Spn SNV*
1	C175	11.33	-	-			
2	C451	8.27	-	-			
3	C453	7.8	-	-			
4	C975	6.03	-	-	p.A338T		p.A1525D
5	W826	4.1	-	-		p.N789T	

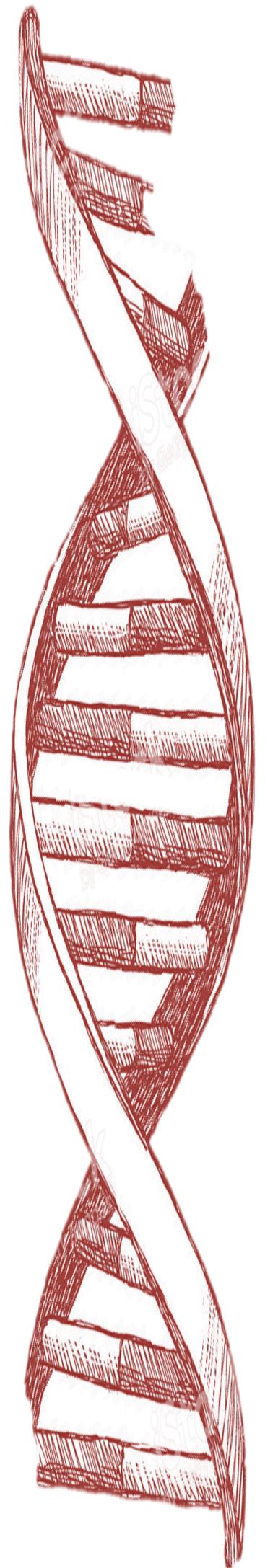
\*: mutaciones confirmadas mediante secuenciación por el método *Sanger*

Por lo tanto, la expresión de *Lmo2* en las HSCs induce leucemia, pero el timo, dónde tiene lugar el segundo evento oncogénico, es indispensable para mantener el fenotipo de T-ALL.





*Discusión*





## DISCUSIÓN

### 1. La pérdida de *Crebbp* a nivel de células pro-B coopera con la sobreexpresión de *Bcl2* para promover la linfomagénesis.

Con el objetivo de estudiar las mutaciones inactivantes de *CREBBP* que se observan en los FL y DLBCL humanos, desarrollamos un modelo de ratón en el que suprimimos la expresión de *Crebbp* en las etapas iniciales de la diferenciación B. Para ello, cruzamos los ratones con uno o ambos alelos de *Crebbp* condicionales<sup>288</sup> con la cepa de ratones *mb1-cre* (Cre recombinasa específica de células pro-B)<sup>289</sup>, con el fin de suprimir la expresión de *Crebbp* en las primeras etapas de la diferenciación B, ya que se ha observado que las mutaciones de *CREBBP* surgen como un evento temprano en el FL<sup>143,185</sup>, lo que nos hizo pensar que las mutaciones de *CREBBP* podrían estar ocurriendo en una etapa inicial de la diferenciación B en estos linfomas, a pesar de que el FL se caracteriza por la presencia de células linfoides tumorales maduras.

La inactivación de *Crebbp* se asoció con reducciones marcadas en las frecuencias de múltiples subtipos de células B de ratones jóvenes y adultos, datos que estaban en línea con observaciones previas en las que la pérdida de *Crebbp*, como consecuencia del uso de la *Cd19-cre*, daba lugar a una reducción del 50% de las células B periféricas<sup>288</sup>. Además, este mismo trabajo mostró que la inactivación combinada de *Crebbp* y *Ep300* daba lugar a la pérdida casi completa de células B<sup>288</sup>. Estas observaciones y nuestros propios datos, por lo tanto, señalan no sólo un importante papel de *Crebbp* en el desarrollo de células B, sino también la redundancia funcional que presenta *Crebbp* con otras histonas acetiltransferasas como *Ep300*.

Por otro lado, varios estudios han descrito la presencia de translocaciones del gen *BCL2* que cooperan con las mutaciones de *CREBBP* tanto en pacientes con linfoma folicular como en aquellos con DLBCL de células B del centro germinal<sup>131, 143</sup>. Además, se piensa que la translocación de *BCL2*, en el linfoma humano, se adquiere a nivel de una célula pre o pro-B<sup>290</sup>. Por ello, teniendo en cuenta nuestras observaciones y con el objetivo de diseñar un modelo que reprodujera lo más fielmente posible la enfermedad humana, decidimos cruzar la cepa *E $\mu$ Bcl2*, que sobreexpresa el oncogén *Bcl2* en las células B de ratón bajo el control del *enhancer* de la cadena pesada de las inmunoglobulinas (IgH)<sup>120</sup>, con los ratones con inactivación de *Crebbp*. Así, observamos que la sobreexpresión de *Bcl2* rescatava parcialmente los déficits en el desarrollo de células B inducidos por la pérdida de *Crebbp*. La cooperación entre la inactivación de *Crebbp* y la sobreexpresión de *Bcl2* también fue demostrada por la

penetrancia y la latencia del desarrollo de linfoma en nuestros modelos transgénicos. Sólo un subconjunto de ratones desarrollaron linfoma en las cepas con inactivación de *Crebbp*, y esto fue más frecuente en los ratones con inactivación bialélica de *Crebbp*, en comparación con los que tenían una inactivación monoalélica, pero su esperanza de vida tampoco fue significativamente diferente a la de los ratones control. Sin embargo, la penetrancia del linfoma se incrementó y la latencia se disminuyó en estos ratones como consecuencia de la adición del transgén *E $\mu$ Bcl2*. La supervivencia específica de linfoma fue significativamente menor en los ratones con una inactivación monoalélica o bialélica de *Crebbp* en combinación con la sobreexpresión de *Bcl2* que en los ratones control *E $\mu$ Bcl2*. Los tumores que desarrollaban estos ratones tenían características de células B de centro germinal, como evidencia el inmunofenotipo B220<sup>+</sup>CD95<sup>+</sup>PNA<sup>+</sup> de las células tumorales y la presencia de expresión de Bcl6 y de hipermutación somática de las inmunoglobulinas en dichas células. Estos datos sugieren que las células B del centro germinal con inactivación de *Crebbp* pueden depender de la coadquisición de uno o más mecanismos para inhibir la apoptosis. A pesar de que nuestros ratones desarrollaban linfomas con una larga latencia, no observamos mutaciones genéticas secundarias que pudieran actuar como eventos secundarios de la linfomagénesis. Sin embargo, la naturaleza clonal de estos tumores sugiere que puede haber algún evento secundario, aún no identificado, que esté impulsando el crecimiento clonal.

Una vez observada la cooperación entre la inactivación de *Crebbp* y la sobreexpresión de *Bcl2* en el desarrollo de linfoma, investigamos las consecuencias moleculares de esta inactivación de *Crebbp* en el contexto de la sobreexpresión de *Bcl2*. Debido a la larga latencia de desarrollo de linfoma en nuestros ratones y con el objetivo de no sesgar nuestros resultados como consecuencia de los efectos de las alteraciones secundarias asociadas a la enfermedad, decidimos llevar a cabo este análisis molecular en una etapa premaligna. Así, realizamos un perfil transcripcional y epigenético de las células B sanas de los ratones que poseían exclusivamente una sobreexpresión de *Bcl2* (cepa control *Cbp<sup>F/F</sup> x E $\mu$ Bcl2*) y de aquellos que además de esta sobreexpresión de *Bcl2* presentaban la inactivación de uno o ambos alelos de *Crebbp* (cepas *Cbp<sup>WT/ $\Delta$</sup>  x E $\mu$ Bcl2* y *Cbp <sup>$\Delta$ / $\Delta$</sup>  x E $\mu$ Bcl2*). El perfil transcripcional de estas células reveló una huella transcripcional alterada como consecuencia de la delección de uno o ambos alelos de *Crebbp*, que incluía la sobreexpresión del oncogén *Myc*. Esta sobreexpresión de *Myc* fue confirmada, además, a nivel de proteína, tanto en los bazos premalignos como en los bazos implicados en linfoma de estos ratones. Sin embargo, nos sorprendió observar que estos cambios transcripcionales no estaban

asociados a alteraciones de la H3K18Ac (acetilación de la H3K18) catalizada por *Crebbp* en los TSSs (sitios de inicio de la transcripción) o en regiones proximales de estos genes. Estos datos sugieren que la desregulación transcripcional observada como consecuencia de la inactivación de *Crebbp* en células B premalignas, podría ocurrir tanto a través de la modificación de elementos reguladores distantes como por una desregulación de la actividad no histona de la proteína CREBBP, o incluso, a través de ambas. Estos datos son consistentes con observaciones previas que mostraban que (1) los cambios epigenéticos impulsados por *Crebbp* presentaban un elevado nivel de redundancia con aquellos que estaban dirigidos por *Ep300*, (2) que estos cambios epigenéticos era dependientes del contexto, y (3) que las dianas no histonas de *Crebbp* podrían ser cruciales en el papel *in vivo* de este gen<sup>307-309</sup>. Además, hemos encontrado evidencias de un enriquecimiento en genes reguladores específicos de células B<sup>310</sup>, entre los genes expresados diferencialmente en las células premalignas de nuestros ratones, pero la confirmación del papel de estos elementos reguladores distantes debe ser llevada a cabo mediante métodos que capturen la conformación tridimensional del genoma, tales como Hi-C<sup>311</sup>.

Por otro lado, observamos cambios significativos en la H3K18Ac de células B con inactivación bialélica de *Crebbp*, los cuales estaban localizados principalmente en regiones intragénicas. Estas regiones con alteración de H3K18Ac mostraban una sobrerrepresentación significativa de secuencias de ADN que poseían sitios de unión a Myc. Además, los datos de ChIP-seq procedentes de la línea celular murina de linfoma de células B (Ch12) y, en menor medida, de la línea celular de eritroleucemia (Mel), mostraron una fuerte unión de Myc al centro de estas regiones con H3K18Ac alterado. Por lo tanto, estos datos sugieren que las alteraciones epigenéticas en estas células B podrían no ser consecuencia directa de la inactivación de *Crebbp*, cuya actividad podría ser compensada por otras acetiltransferasas<sup>288</sup>, pudiendo ser consecuencia del aumento de la expresión de *Myc* y de su actividad en regiones intragénicas potenciadoras. Es importante destacar que se ha demostrado que el papel de *Myc* en estos elementos “superpotenciadores” es crítico para la etiología del DLBCL<sup>312</sup>.

## 2. Presencia de diferentes patrones mutacionales de *CREBBP* en el linfoma folicular en comparación con el linfoma difuso de células B grandes.

Nuestro modelo murino pone en evidencia importantes contrastes con el FL, el cual tiene una mayor frecuencia de mutaciones de *CREBBP* que el DLBCL<sup>131,132</sup>. En primer lugar, la mayoría de los linfomas desarrollados por nuestros ratones transgénicos parecen más similares a un FL de alto grado o a un DLBCL. Sin embargo, las mutaciones de *CREBBP* son más frecuentes en el FL humano de bajo grado<sup>143</sup>. En segundo lugar, la delección de *Crebbp* en nuestro modelo murino se asoció con la inducción de la expresión de *Myc*, sobreexpresión que rara vez se describe en el FL de bajo grado<sup>313</sup>, pero que sí es más frecuente en el DLBCL<sup>314,315</sup>. Estas diferencias nos llevaron a examinar si la mutación de *CREBBP* podía tener papeles divergentes en estas dos enfermedades. Para ello, caracterizamos las mutaciones de *CREBBP* en un gran número de tumores de pacientes con FL y DLBCL, identificando un espectro significativamente diferente de mutaciones de *CREBBP* en estos dos tipos de linfoma. Este análisis nos permitió comprobar que los DLBCL presentaban más frecuentemente mutaciones de cambio del marco de lectura (*nonsense/frameshift*) que ocurrían aguas arriba del dominio KAT y que probablemente tenían como resultado la pérdida de la proteína, dando lugar a un escenario comparable al desarrollado en nuestro modelo murino con delección de *Crebbp*. Por el contrario, los FL presentaban con mayor frecuencia mutaciones de sustitución de un único aminoácido (*missense/in-frame*) en el dominio KAT, que probablemente producían una proteína catalíticamente inactiva. Por ello, sugerimos que estos dos patrones mutacionales probablemente tienen consecuencias funcionales diferentes. En línea con esta observación, sólo encontramos un modesto cambio en la expresión del MHC-II de las células B de los ratones con delección de *Crebbp*, mientras que un análisis previo del FL humano reveló una marcada desregulación de la expresión del MHC-II asociada a mutaciones puntuales de *CREBBP*<sup>143</sup>.

Además, no se observó ningún cambio significativo en la expresión de *MYC* asociado a mutaciones de *CREBBP* en el FL humano, pero sí se encontraron evidencias de una mayor expresión de *MYC* asociada a mutaciones de *CREBBP* en el DLBCL humano. Un apoyo adicional de que la mutación de *CREBBP* está asociada con la inducción de la expresión de *MYC* en los DLBCL, la observamos en la exclusividad mutua entre las mutaciones de *CREBBP* y las translocaciones de *MYC*. Esto dibuja paralelismos entre nuestras observaciones en el modelo murino con delección de *Crebbp* y el DLBCL humano, y proporciona evidencias de un posible papel potencial de las mutaciones de

*CREBBP* en la inducción de la expresión de *MYC* en esta enfermedad. Además, es importante destacar que *CREBBP* puede controlar la expresión de *MYC* a través de su inhibición, por la vía de la  $\beta$ -*catenina*, mediante la acetilación específica de la lisina 49<sup>316</sup>. La interferencia de esta acetilación, como consecuencia de una mutación inactivante de *CREBBP* o por la mutación de la lisina 49 de la  $\beta$ -*catenina*, lleva a la transactivación específica de *MYC*, pero no de otros genes regulados por esta vía<sup>316</sup>.

Estos datos proporcionan un potencial vínculo funcional entre la pérdida de *CREBBP* y la activación de *MYC*. Sin embargo, el papel de la mutación de *CREBBP* en la inducción de *MYC* en el DLBCL humano, así como, la divergencia funcional de las mutaciones de *CREBBP*, son áreas que requieren de nuevos análisis y validaciones en experimentos futuros. Así, para poder proporcionar una firme asociación entre la inactivación de *CREBBP* y la inducción de *MYC* en el DLBCL humano, una amplia cohorte de tumores tendrían que ser analizados mediante inmunohistoquímica para *MYC* y, en paralelo, se deberían llevar a cabo estudios en dichos tumores de las mutaciones de *CREBBP*, FISH para la translocación t(8;14) asociada a *MYC*, y el análisis de otras alteraciones genéticas que podrían estar afectando a la expresión de *MYC*, tales como la ganancia en el número de copias del ADN<sup>317</sup> o la mutación de *EZH2*<sup>318</sup>.

En conclusión, nuestros resultados demuestran que *Crebbp* tiene un papel crucial en el desarrollo de células B, y la pérdida de *Crebbp* a nivel de células pro-B colabora con la sobreexpresión de *Bcl2* para promover el DLBCL de células B de centro germinal. Además, estos linfomas tienen perfiles moleculares que se caracterizan por la sobreexpresión de *Myc* y alteraciones epigenéticas de regiones intragénicas unidas a *Myc*. Por otro lado, nuestros datos ponen de relieve diferencias importantes entre las mutaciones de *CREBBP* de los DLBCL humanos en comparación con las que aparecen en los FL humanos, y sugieren que las mutaciones sin sentido inducidas por un cambio del marco de lectura (*nonsense/frameshift*) de *CREBBP*, podrían tener un papel importante en la inducción de *MYC* en el DLBCL.

### 3. La expresión transitoria de *Lmo2* en células madre hematopoyéticas (HSCs) causa T-ALL dependiente del timo.

El estudio de los eventos escalonados que tienen lugar durante la evolución de las células tumorales es difícil debido, en primer lugar, a las numerosas alteraciones genéticas que se vuelven clonalmente dominantes en el momento de su detección dentro del tumor clínicamente manifestado, y en segundo lugar, debido a la multitud de vías por las cuales un tumor puede evolucionar<sup>319</sup>. En principio, la leucemogénesis es un proceso mediante el cual una célula normal adquiere una identidad nueva, pero inapropiada (maligna), que da lugar a una población clonal aberrante. Sin embargo, este hecho sólo es posible si la célula normal que origina la leucemia, es decir, la célula de origen de la leucemia, tiene la plasticidad necesaria para experimentar cambios. Además, los eventos oncogénicos que inician el cáncer deben tener una capacidad de reprogramación inherente, para poder inducir el cambio en la identidad celular de la célula de origen<sup>320</sup>. A pesar de que estudios previos han demostrado el papel de diferentes alteraciones genéticas en la iniciación de algunas neoplasias hematológicas, a través de un mecanismo de reprogramación tumoral de las HSC/PCs<sup>18-23</sup>, aún no se ha sugerido que un mecanismo similar pueda estar ocurriendo en las T-ALL.

Para estudiar si efectivamente este mecanismo de reprogramación tumoral está dirigiendo las T-ALL asociadas a terapia génica, nos planteamos generar un modelo de ratón condicional de *Lmo2* (cepa *Rosa26-Lmo2+Sca1-Cre*) con el fin de analizar su capacidad para reprogramar las HSCs hacia un fenotipo maligno. Así, nuestros resultados proporcionan la primera evidencia genética *in vivo* de que la expresión de *Lmo2* en HSCs, por sí sola, es capaz de dirigir la diferenciación de las HSC/PCs hacia un fenotipo tumoral que recapitula el inmunofenotipo y las características moleculares y transcripcionales de la T-ALL humana *LMO2+*. Por lo tanto, estos resultados indican que la célula de origen de las T-ALL debe poseer suficiente plasticidad para permitir que se lleve a cabo la reprogramación tumoral (o al menos, para que se inicie), y las características finales de la leucemia serán el resultado de la interacción entre la capacidad de reprogramación del oncogén *Lmo2* y la susceptibilidad de la célula de origen (HSC) a dicha reprogramación.

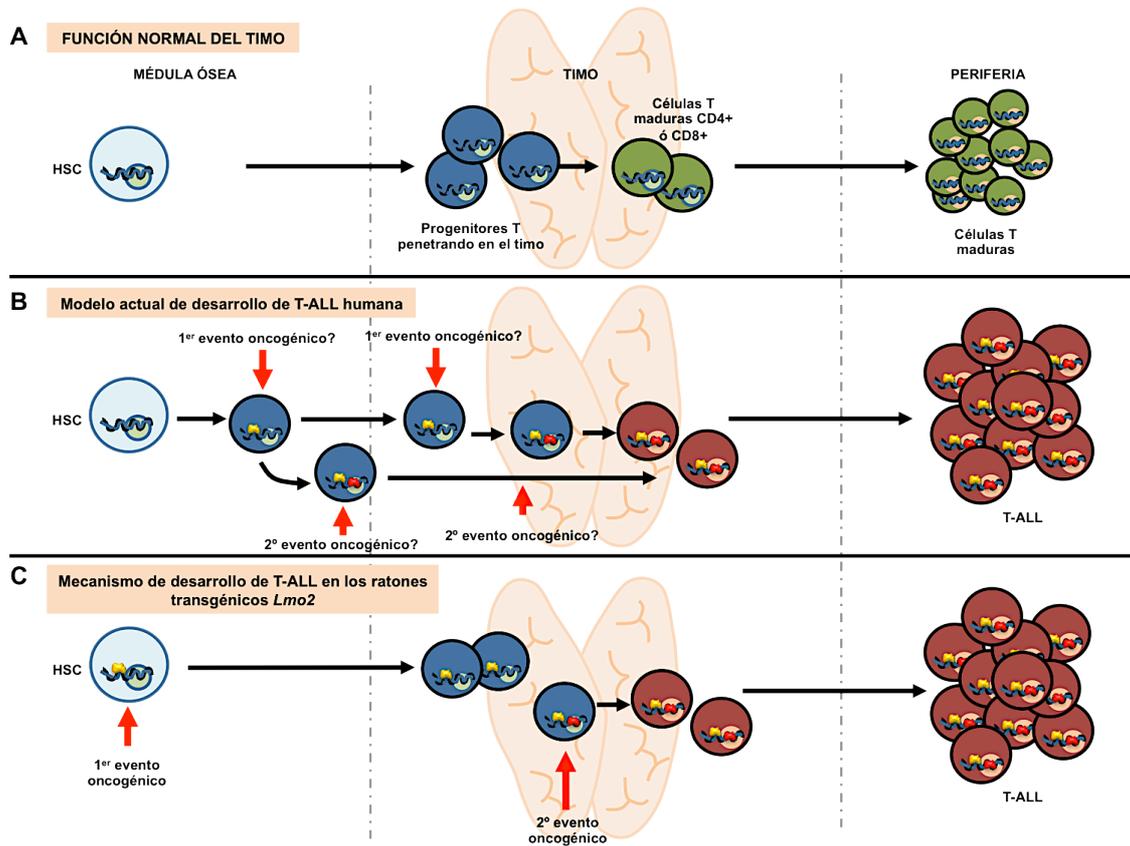
El hecho de que *Lmo2* esté implicado en la reprogramación no patológica de células diferenciadas hacia HSCs<sup>259</sup>, junto con nuestros datos, nos llevó a hipotetizar que *Lmo2* podía estar actuando mediante un mecanismo "*hit-and-run*", promoviendo la oncogénesis a través de su expresión transitoria durante una etapa temprana de la

evolución tumoral y dirigiendo la diferenciación de la célula de origen (HSC) hacia un fenotipo T maligno. Por ello, intentamos demostrar esta hipótesis mediante la expresión transitoria de *Lmo2* en HSC/PCs murinas (cepa *Sca1-Lmo2*). Las T-ALL que desarrollaron estos ratones fueron histológica, transcripcional y genéticamente similares a la T-ALL humana dirigida por *LMO2*. Sin embargo, las células T de la masa tumoral no mostraron expresión de *Lmo2*, evidenciando que la expresión del mismo no era necesaria para el mantenimiento del fenotipo de T-ALL. Estos datos muestran que sólo aquellos progenitores T que pierden la expresión de *Lmo2* son seleccionados clonalmente para su posterior transformación tumoral. Estos datos no sólo resaltan un papel previamente desconocido para *Lmo2* en esta enfermedad, sino que también proporcionan la evidencia de un proceso previamente no modelado para las T-ALL, donde la programación de la identidad tumoral ya se ha realizado en la etapa de la célula de origen.

Dada la alta frecuencia de aparición de mutaciones en importantes vías oncogénicas (*Notch1* y *Ras*) en las T-ALL desarrolladas por nuestros ratones, nos planteamos determinar dónde estaban teniendo lugar (en la médula ósea o en el timo) estos eventos secundarios. Para ello, diseñamos un modelo murino *Sca1-Lmo2* carente de timo (cepa *Sca1-Lmo2<sup>nu/nu</sup>*). Las leucemias desarrolladas por estos ratones se caracterizaban por ser dependientes de la expresión de *Lmo2* y, además, no mostraban mutaciones recurrentes ni de *Notch1* ni de *Ras*. Estos datos, por lo tanto, demuestran que la expresión de *Lmo2* en las HSCs induce leucemia, pero el timo, dónde tiene lugar el segundo evento oncogénico, es indispensable para establecer el fenotipo de T-ALL.

La implicación de estos hallazgos es que, dado que la T-ALL puede desarrollarse a través de un mecanismo de diferenciación dirigida similar a la reprogramación, los oncogenes que inician la formación de tumores pueden ser prescindibles para la supervivencia de las células tumorales y/o la progresión tumoral. En este contexto, las mutaciones que activan a los oncogenes tendrían un papel impulsor en el proceso de reprogramación, pero pueden actuar como mutaciones pasajeras a partir de entonces, o pueden tener un papel secundario en los clones de células tumorales evolucionadas. Basándonos en nuestras observaciones, proponemos un modelo de "*hit-and-run*" en el que el oncogén *Lmo2* puede actuar sobre los precursores de las células tumorales tempranas, no siendo requerido en la progenie de las células tumorales desarrolladas (**Figura 42**). Las células tumorales desarrolladas pueden, a su vez, volverse dependientes de vías de supervivencia alternativas que no están presentes en sus

precursores. Por ello, proponemos que en la T-ALL humana, las alteraciones genéticas de *LMO2* pueden actuar de manera fulminante en los primeros precursores, mientras que las células tumorales evolucionadas desarrollan dependencia de mecanismos oncogénicos alternativos, como *Notch1* o *Ras*.



**Figura 42: Modelo propuesto por el cual la expresión ectópica de *Lmo2* reprograma a las HSC/PCs hacia un fenotipo tumoral de células T.** **A)** Desarrollo linfóide normal de células T en humanos y ratón. Los círculos verdes representan los eventos genéticos reguladores normales (activadores o represores) que ocurren durante el desarrollo temprano de la diferenciación T. Los círculos salmón representan los eventos genéticos reguladores que ocurren durante la maduración final de células T. **B)** Modelo actual de desarrollo de la T-ALL humana. Se desconoce la existencia de alteraciones latentes previas a la diferenciación terminal. Por lo tanto, se desconoce la naturaleza de la célula de origen cancerígena y la localización celular donde se produce el segundo evento oncogénico. **C)** Mecanismo de desarrollo de T-ALL en los ratones *Sca1-Lmo2*. Los semicírculos amarillos representan eventos epigenéticos reguladores latentes causados por la expresión de *Lmo2* en HSC. Estas marcas epigenéticas no interfieren con el desarrollo normal de las células T, pero se vuelven activas (ya sea activando o reprimiendo) en el proceso de diferenciación terminal cuando aparece un segundo evento dentro del timo (semicírculos rojos), lo que conduce a la aparición de las células T tumorales. Según este modelo, la T-ALL es el resultado de un proceso de reprogramación celular que se inicia en una HSC.

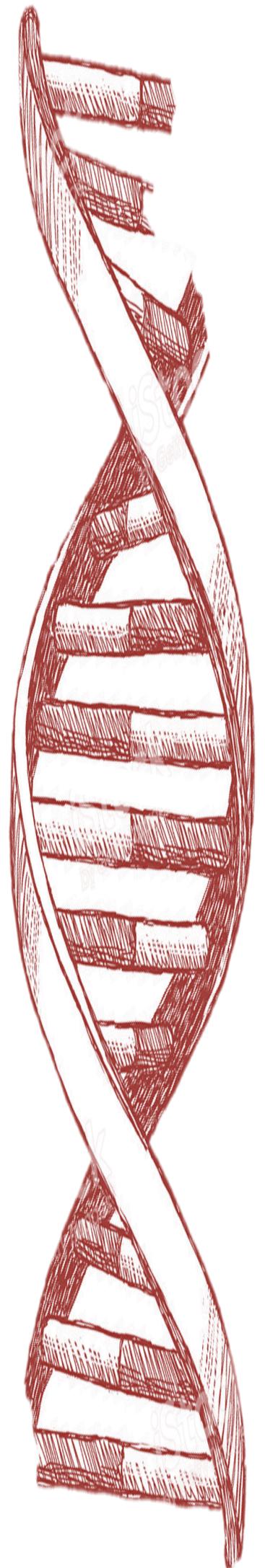
De hecho, nuestros resultados demuestran que la presencia del oncogén es necesaria en las primeras etapas de la transformación, pero el fenotipo final del tumor está determinado por el nicho del timo. Esto puede proporcionar una explicación para el fracaso de algunas terapias modernas para eliminar las células madre tumorales, a pesar de ser agentes efectivos contra las células tumorales evolucionadas. Consecuentemente, estrategias de tratamiento dirigidas tanto a las vías oncogénicas que son activas en las etapas tempranas, como a aquellas que son activas en etapas tardías del desarrollo del tumor, pueden ser necesarias para erradicar por completo las T-ALL. Por lo tanto, nuestros hallazgos tienen importantes implicaciones para comprender y atacar terapéuticamente las células tumorales de las T-ALL.

En el contexto de la terapia génica *ex vivo*, nuestros datos muestran claramente que *LMO2* es el evento oncogénico iniciador que induce la reprogramación tumoral de las HSCs hacia el fenotipo de T-ALL. En el futuro, la extinción de la transactivación de *Lmo2* en las HSC utilizando inactivadores gamma-retrovirales<sup>321</sup> o vectores lentivirales<sup>322</sup> evitará el desarrollo de T-ALL en pacientes con SCID-X1 tratados con terapia génica *ex vivo*, contribuyendo al desarrollo de ensayos clínicos más seguros.





*Conclusiones*





### CONCLUSIONES

#### PRIMERA

- *Crebbp* tiene un papel crucial en la génesis del linfoma difuso de células B grandes de centro germinal, sugiriendo que es el evento genético, y no la célula de origen, quien condiciona la identidad tumoral.
- Además, nuestros datos ponen de relieve las importantes diferencias que existen entre las mutaciones de *CREBBP* en los linfomas difusos de células B grandes humanos en comparación con las mutaciones que aparecen en los linfomas foliculares humanos, sugiriendo que las mutaciones de cambio del marco de lectura (*frameshift/nonsense*) de *CREBBP* tienen un papel importante en la inducción de *MYC* en el linfoma difuso de células B grandes.

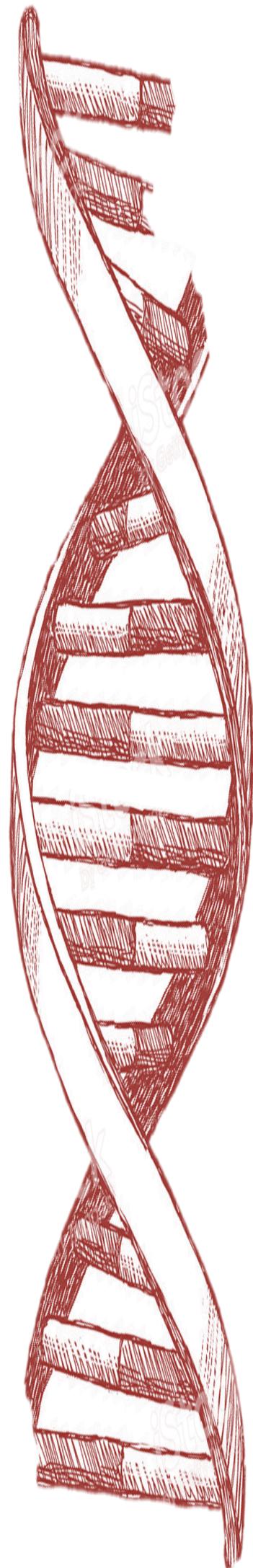
#### SEGUNDA

- La expresión transitoria de *Lmo2* en el compartimento de células madre/progenitoras hematopoyéticas (HSC/PCs) es capaz de inducir una leucemia linfoblástica aguda de células T similar a la humana, como consecuencia de un mecanismo de reprogramación tumoral. Estos resultados, igualmente, ponen de manifiesto que es el evento oncogénico quien define la identidad tumoral.
- Asimismo, nuestros datos demuestran que las alteraciones genéticas secundarias (mutaciones de *Notch1* y *Ras*) de la leucemia linfoblástica aguda de células T tienen lugar en el timo.





*Material y métodos*





## MATERIAL Y MÉTODOS

### 1. Animales utilizados en el estudio.

La generación de los ratones *mb1-cre*<sup>289</sup>, *EμBcl2*<sup>120</sup> y los ratones *Crebbp*<sup>WT/F</sup> y *Crebbp*<sup>F/F</sup><sup>288</sup> ha sido descrita previamente. Los modelos animales *Cbp*<sup>WT/Δ</sup> y *Cbp*<sup>Δ/Δ</sup> fueron generados mediante el cruce de las cepas *Crebbp*<sup>WT/F</sup> y *Crebbp*<sup>F/F</sup> con la cepa *mb1-cre*. Los modelos animales *Cbp*<sup>WT/Δ</sup> x *EμBcl2* y *Cbp*<sup>Δ/Δ</sup> x *EμBcl2* fueron generados gracias al cruce de las cepas *Cbp*<sup>WT/Δ</sup> y *Cbp*<sup>Δ/Δ</sup> con los ratones *EμBcl2*.

Los ratones *Rosa26-Lmo2* fueron generados de la siguiente manera: el ADNc murino de *Lmo2* fue clonado dentro del plásmido Rosa26UA usando el sitio *Ascl*. Este plásmido contenía la construcción *Loxp-pgk-Neo-tPA-Loxp-Ascl-FheI-IRES-GFP* dentro de los brazos de homología del locus Rosa26, obteniendo la construcción final *Loxp-pgk-Neo-tPA-Loxp-Lmo2-IRES-GFP*. El vector fue electroporado en células ES (embryonic stem o madre embrionarias) de ratón para su recombinación homóloga y se aislaron aquellos clones de células ES resistentes a neomicina. Se seleccionaron los recombinantes homólogos mediante PCR, análisis de *Southern Blot* y cariotipo, y se inyectaron en la masa celular interna de blastocistos de embriones de ratón. Los embriones se transfirieron al útero de una hembra pseudopreñada y se generaron quimeras. Éstas se cruzaron con ratones C57BL/6 para obtener heterocigotos. Se estableció, así, la línea de ratones *Rosa26-Lmo2*. A continuación, dichos ratones se cruzaron con ratones *Sca1-Cre*<sup>293</sup> para generar la cepa de ratones *Rosa26-Lmo2+Sca1-Cre*. Las microinyecciones se realizaron en colaboración con la Dra. Belén Pintado del servicio de transgénesis del Centro Nacional de Biotecnología de Madrid (CNB).

La generación de los ratones *Sca1-Lmo2* ha sido descrita previamente<sup>294</sup>. Los ratones *Sca1-Lmo2*<sup>nu/nu</sup> se obtuvieron mediante el cruce de la cepa *Sca1-Lmo2* con ratones desnudos *nu/nu*.

Todos los ratones utilizados en este trabajo tienen un *background* mixto CBAxC57BL/6J. Para el genotipado de los ratones se usaron las técnicas de *Southern-blot* y PCR, que se explican en detalle en los siguientes apartados.

El trabajo con animales ha sido realizado siguiendo las pertinentes directrices nacionales e internacionales, y ha sido aprobado por el Comité de Bioética de la Universidad de Salamanca (USAL) y por el Subcomité de Bioética del Consejo

Superior de Investigaciones Científicas (CSIC). Para la realización de este proyecto he obtenido la capacitación para experimentadores en ciencias del animal de laboratorio, Categoría B.

Cuando fueron observados signos de enfermedad, los animales fueron sacrificados y se les realizaron procedimientos estándar de necropsia. Los análisis de supervivencia fueron realizados utilizando el test estadístico *log-rank* (Mantel-Cox).

### 2. Genotipado de los ratones mediante *Southern blot*.

La cepa de ratones *Rosa26-Lmo2* se genotipó mediante la técnica de *Southern Blot*. Para ello, el ADN genómico de los ratones se obtuvo a partir de un fragmento de cola siguiendo el procedimiento que se describe en la sección “8. Extracción de ADN”. El ADN se resuspendió en 50µL-70µL de MPW. 10 µg de ADN se incubaron a 37°C con 20U de la endonucleasa de restricción EcoRI. Tras 5 horas de digestión, la reacción fue detenida y los fragmentos obtenidos, separados mediante electroforesis en un gel de agarosa. Los fragmentos de ADN separados mediante electroforesis se desnaturalizaron mediante la inmersión del gel en una solución alcalina desnaturalizante (1,5M NaCl, 0,5M NaOH) y agitación continua a temperatura ambiente durante 45 min, con el objetivo de romper los enlaces de hidrógeno que unen las dos cadenas y obtener fragmentos de ADN monocatenario. Acto seguido, se neutralizó la reacción de desnaturalización con una solución neutralizante (1,5M NaCl; 1M Tris- HCl pH 7,5) durante 45 min a temperatura ambiente y agitación continuada. A continuación, se realizó la transferencia del ADN a una membrana de nylon (Ref. RPN 303-N, GE Healthcare) durante 16 horas, utilizando como tampón de transferencia, 20X SSC (6M NaCl, 0,6M citrato sódico dibásico, pH 7.0). Tras la transferencia, y con el fin de conseguir una unión covalente del ADN a la membrana, ésta fue sometida a una fuente de luz UV de 312nm durante 20 s. (UV Stratalinker 2400, Stratagene). Por último, se llevo a cabo la hibridación del ADN unido a la membrana. Para ello, usamos como sonda radiactiva el fragmento “RA-4”, procedente de un plásmido del laboratorio de *Stuart Orkin* en Boston. Se trata de un vector que contiene 4 kb (kilobases) de ADN genómico del locus *Rosa26* y, mediante digestión con EcoRI y PstI, se obtuvo un fragmento de 693 pb (pares de bases) que se marcó radiactivamente con [ $\alpha^{32}$ P] dCTP y se usó como sonda para hibridar con el ADN genómico unido a la membrana. En paralelo, se aplicaron técnicas de reacción en cadena de la polimerasa (PCR), que describimos en detalle en el siguiente apartado.

Los ratones *Sca1-Cre* se genotiparon mediante *Southern Blot* siguiendo el procedimiento descrito anteriormente y usando la sonda “Cre” que corresponde a

parte del ADNc del gen codificante de la Cre recombinasa. Ésta se hibridó con el ADN genómico del ratón digerido con 20U de la endonucleasa de restricción EcoRI. El genotipado de los ratones *Sca1-Lmo2* se ha descrito previamente<sup>294</sup>.

### 3. Genotipado de los ratones mediante PCR.

Para el genotipado de algunas cepas se llevó a cabo PCR del ADN total extraído de un fragmento de cola de los ratones, siguiendo el procedimiento que se describe en detalle en la sección “8. Extracción de ADN”.

#### 3.1 PCR de GFP.

La cepa de ratones *Rosa26-Lmo2* fue genotipada, también, mediante PCR de *GFP* y *neo*. Para la PCR de *GFP* se preparó una mezcla de reacción conteniendo 2,5µL del oligo de sentido 10µM; 2,5µL del oligo anti-sentido 10µM; 2,5µL de un *buffer* comercial (Ref. 2900162, 5PRIME); 2,5µL dNTPs 2mM; 2,5µL MgCl<sub>2</sub> 25mM (Ref. 2900154, 5PRIME); 1U del enzima Taq DNA polimerasa (Ref. 2900163, 5PRIME); 11,3µL MPW, 1µL de ADN y una gota de aceite mineral (Ref. M3516-6ML, SIGMA). Se usaron los siguientes oligos:

GFP-FP2                      sentido: 5´ AGC ACG ACT TCT TCA AGT CC 3´

GFP-RP1                      anti-sentido : 5´ TCC ATG CCG AGA GTG ATC C 3´

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) mediante el siguiente programa: 95°C, 4´+ [94°C, 30´+ 58°C, 30´+ 72°C, 1´] x 35 ciclos + 72°C, 10´ + 4°C . El producto de la amplificación tenía un tamaño de 450 pb y se corrió en un gel de agarosa al 2% a 65V durante 1 hora.

#### 3.2 PCR de neo.

Para la PCR de *neo* de los ratones *Rosa26-Lmo2* se preparó una mezcla de reacción conteniendo 2,5µL del oligo “oIMR013” 10µM; 2,5µL del oligo “oIMR014” 10µM; 2,5µL del oligo “oIMR015” 10µM; 2,5µL del oligo “oIMR016” 10µM; 5µL de un *buffer* comercial Green GoTaq Flexi *Buffer* (Ref. M891A GoTaqFlexi DNA Polymerase, Promega); 2,5µL dNTPs 2mM; 2,5µL MgCl<sub>2</sub> 15mM; 1U del enzima GoTaq DNA Polymerase (Ref. M829A, Promega); 3,8µL MPW, 1µL de ADN y una gota de aceite mineral (Ref. M3516-6ML, SIGMA). Se usaron los siguientes oligos:

oIMR013: 5´ CTT GGG TGG AGA GGC TAT TC 3´



Oligo mb-1 sen: 5' CTG CGG GTA GAA GGG GGT C 3'

Oligo mb-1 antisen: 5' CCT TGC GAG GTC AGG GAG CC 3'

Oligo hCre sen: 5' CCC TGT GGA TGC CAC CTC 3'

Oligo hCre antisen: 5' GTC CTG GCA TCT GTC AGA G 3'

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) mediante el siguiente programa: 95°C, 4'30''+ [95°C, 30''+ 65°C, 45''+ 72°C, 30''] x 36 ciclos + 72°C, 2' + 4°C. El producto de la amplificación usando los oligos "mb1" tenía un tamaño de 420 pb, mientras que el resultante del uso de los oligonucleótidos "hCre" tenía uno de 430 pb. Ambos productos se cargaron en un gel de agarosa al 2% y se realizó la electroforesis utilizando una diferencia de potencial de 65V durante una hora. Sólo la presencia de las dos bandas en el gel indicaba un genotipo *mb1-Cre*. La presencia de una única banda de 420 pb determinaba un genotipo silvestre (*WT*).

### **3.5 PCR *Emu-Bcl2*.**

Los ratones *EmuBcl2* fueron genotipados mediante PCR de *Emu-Bcl2*. Se preparó una mezcla de reacción conteniendo 2,5µL del oligo de sentido 10µM; 2,5µL del oligo anti-sentido 10µM; 2,5µL de un *buffer* comercial (Ref. 2900162, 5PRIME); 2,5µL dNTPs 2mM; 2,5µL MgCl<sub>2</sub> 25mM (Ref. 2900154, 5PRIME); 1U del enzima Taq DNA polimerasa (Ref. 2900163, 5PRIME); 11,3µL MPW, 1µL de ADN y una gota de aceite mineral (Ref. M3516-6ML, SIGMA). Se usaron los siguiente oligos:

Emu-Bcl2                      sentido: 5' GCA GAC ACT CTA TGC CTG TGT GG 3'

Emu-Bcl2                      anti-sentido: 5' GGA ACC TTA CTT CTG TGG TGT GA 3'

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) mediante el siguiente programa: 94°C, 4'+ [94°C, 45''+ 60°C, 45'+ 72°C, 45''] x 35 ciclos + 72°C, 5' + 4°C. El producto de la amplificación tenía un tamaño de 500 pb y se corrió en un gel de agarosa al 2% a 65V durante 1 hora.

### 4. Análisis de citometría de flujo.

Se obtuvieron leucocitos de médula ósea total (haciendo *flushing* de los huesos largos), sangre periférica, timo y bazo de los ratones. Los eritrocitos fueron eliminados mediante la lisis de los mismos con solución de lisis (RCLB, por sus siglas en inglés: *Red Cell Lysis Buffer*) y las células restantes fueron lavadas con PBS con 1% de FCS (PBS-1%FCS). Tras ser teñidas, las células se lavaron una vez más con PBS-1%FCS y posteriormente teñidas con yoduro de propidio (PI) a una concentración de 2 mg/ml para permitir que las células muertas fueran excluidas tanto del análisis como de los procedimientos de clasificación. Las muestras fueron adquiridas en un citómetro ACCURI C6 y analizadas utilizando el software informático Flowjo. La fluorescencia específica para los fluorocromos FITC, PE, PI, PerCP, PE-Cy7 y APC, excitados a 488nm (0,4W) y 633nm (30mW), así como las propiedades conocidas de dispersión de la luz directa y ortogonal de las células de ratón fueron utilizadas para establecer los *gates*. En cada análisis fue examinado un mínimo de 50.000 células viables (negativas para PI).

Los siguientes anticuerpos fueron utilizados en los análisis de citometría de flujo: anti-B220 (RA3-6B2; 1:100), anti-CD3 $\epsilon$  (145-2C11; 1:100), anti-CD4 (RM4-5; 1:250), anti-CD8a (53-6.7; 1:250), anti-CD11b/Mac1 (M1/70; 1:200), anti-CD19 (1D3; 1:100), anti-CD117/c-Kit (2B8; 1:200), anti-Ly-6G/Gr1 (RB6-8C5; 1:100), anti-IgD (11-26c.2a; 1:100), anti-IgM (R6-60.2; 1:100), anti-Sca1/Ly-6A/E (E13-161.7; 1:50), anti-CD21 (7G6; 1:100), anti-CD23 (B3B4; 1:100), anti-CD25 (PC61; 1:500), anti-I-A/I-E (2G9; 1:100), anti-IgG (Poly4053; 1:250), anti-CD95 (Jo2; 1:100), anti-PNA (FL-1071; 1:100), anti-Ter119 (TER119, 1:100). La posible unión inespecífica de los anticuerpos utilizados fue suprimida mediante la pre-incubación previa de las muestras con anti-CD16/CD32 Fc-block (2.4G2; 1:100). Todos los anticuerpos fueron adquiridos de BD Biosciences.

### 5. Estudio de la formación del centro germinal.

Para estudiar la formación de centros germinales en los ratones *Crebbp*, se procedió a la inmunización de los ratones *Crebbp*<sup>WT/F</sup>, *E $\mu$ Bcl2*, *Cbp*<sup>WT/ $\Delta$</sup> , *Cbp* <sup>$\Delta/\Delta$</sup> , *Cbp*<sup>WT/ $\Delta$</sup>  x *E $\mu$ Bcl2* y *Cbp* <sup>$\Delta/\Delta$</sup>  x *E $\mu$ Bcl2* con una solución de eritrocitos de cordero (SRBCs, *sheep red blood cells*, Ref. SR0051C, Thermo Scientific). Se tomaron 5mL de la solución concentrada y se diluyeron en 50mL de PBS (tampón fosfato salino) 1X Na<sup>+</sup>. Se procedió a una centrifugación de 10 minutos a 2500 rpm y el botón celular se volvió a resuspender en 50mL de PBS 1X Na<sup>+</sup>. Estos dos últimos pasos se repitieron 3 veces más y la solución se llevó finalmente a una concentración de 2x10<sup>9</sup> células/mL. Una vez preparada la

solución, a la concentración indicada, se inyectaron 100  $\mu\text{L}$  de la misma ( $1-2 \times 10^8$  células) intraperitonealmente en todas las cepas. El tiempo óptimo para analizar la reacción del centro germinal es de 10-12 días<sup>323</sup>, período tras el cual, se sacrificaron los animales inmunizados. Se aisló el bazo de dichos ratones, tomándose parte de él para su análisis por anatomía patológica y, otra parte, para su estudio por citometría de flujo. Específicamente, se usaron 200  $\mu\text{L}$  de bazo ( $2 \times 10^6$  células) y éste se tiñó con los anticuerpos B220, PNA y CD95, para determinar, así, la capacidad de formación de células B del centro germinal.

### **6. Análisis histológicos e inmunohistoquímica (IHC).**

Los animales fueron sacrificados mediante dislocación cervical y las muestras obtenidas de los mismos fueron fijadas en formaldehído 4% e incluidas en parafina. Los estudios patológicos fueron realizados sobre secciones teñidas con hematoxilina-eosina bajo la supervisión del experto en patología Óscar Blanco, del hospital Universitario de Salamanca, y de expertos hematopatólogos de la Universidad de Nebraska.

Para llevar a cabo las tinciones inmunohistoquímicas, las muestras de los ratones transgénicos fueron cortadas, desparafinadas y calentadas en citrato sódico 10 mmol/L, durante 30 minutos. Los cortes fueron incubados con anticuerpos primarios. Los anticuerpos usados incluyeron: anti-Pax5 (Clon-1EW, prediluido, Leica Biosystems), anti-MYC (Clon-Y69, prediluido, Master Diagnostica), anti-Bcl6 (Clon-LN22, prediluido, Leica Biosystems), anti-PAX5 (clon 24/PAX5, 1:100, BD Biosciences), anti-TdT (1:50, Supertechs INC) y anti-CD3 (clon SP7, 1:50, Abcam).

### **7. Análisis de expresión de proteína mediante *Western blot*.**

Se realizaron estudios de *Western blot* en células T tumorales del timo de ratones *Rosa26-Lmo2+Sca1-Cre* leucémicos, para analizar la expresión de la proteína LMO2 en los ratones leucémicos. El timo sano de ratones *control WT* fue usado como control negativo.

Las células del timo fueron lisadas con *buffer* RIPA (50mM Tris pH 8.0, 150mM NaCl, 0.5% desoxicolato de sodio, 1% NP-40, 0.1% SDS) que contenía proteasa e inhibidores de fosfatasas (Roche Diagnostics), obteniendo así la proteína total de cada muestra. 50  $\mu\text{g}$  de proteína total fueron separados en SDS-PAGE y transferidos a membranas Hybond-C Extra (Amersham Biosciences).

La inmunotransferencia fue llevada a cabo utilizando los siguientes anticuerpos: anti-LMO2 1:500 (Abcam, #ab72841) y anti  $\beta$ -Actina 1:7500 (Santa Cruz, clon I-19,

#sc1616). La detección fue llevada a cabo utilizando anticuerpos anti-conejo (BioRad, Ref.70-6515) y anti-cabra (Santa Cruz, Ref. sc2033) conjugados con peroxidasa de rábano (HRP) mediante un sistema de detección ECL (Thermo Scientific, Ref. 80196).

### 8. Extracción de ADN

Los tejidos fueron disgregados con 500µl o 5ml, en función de la celularidad de cada tejido, de tampón proteinasa K (0.25M de sacarosa, 50mM Tris pH 7.4, 25mM KCl, 5mM MgCl<sub>2</sub>; 1% SDS; 20mM EDTA pH 8; 1µg de proteinasa K (Roche)) durante 16h a 55°C. Tras este periodo de incubación, se realizó una extracción fenol-cloroformo para limpiar el ADN de restos celulares. Brevemente, se añadió a la muestra incubada un volumen de fenol (Gibco BRL), se centrifugó 5 minutos a máxima rpm y temperatura ambiente y nos quedamos con la fase superior. Después, se añadió la fase superior en medio volumen de fenol (Gibco BRL) y medio volumen de cloroformo (Merk), realizándose otra centrifugación similar a la anterior, y quedándonos con la fase superior. Posteriormente, se añadió la fase superior a un volumen de cloroformo (Merk), realizándose el mismo protocolo de centrifugación. La fase superior obtenida se precipitó en 2 volúmenes de etanol 100% (Merk). El precipitado de ADN obtenido se lavó con etanol 70%. Posteriormente, se resuspendió el ADN en agua bidestilada (ddH<sub>2</sub>O), cuantificándose la concentración de ADN mediante la medida de absorción a 260nm con el espectrofotómetro Nanodrop-1000 (THERMO Scientific).

### 9. Análisis de clonalidad de las inmunoglobulinas (VDJ).

Los reordenamientos de las inmunoglobulinas fueron amplificados por PCR utilizando los oligos listados al final de esta sección. Las condiciones de la PCR consistieron en una activación inicial a 95°C seguida por 31-37 ciclos de desnaturalización (1 minuto a 95°C), anillamiento (1 minuto a 65°C), y elongación (1 minuto y 45 segundos a 72°C). Este paso fue seguido por una elongación final de 10 minutos a 72°C. Para determinar las secuencias de ADN de cada reordenamiento V(D)J individual, los fragmentos de PCR fueron aislados del gel de agarosa 0,8% y clonados en el vector pGEM-Teasy (Promega); Los insertos de ADN de, al menos, 10 clones correspondientes al mismo fragmento de la PCR, fueron secuenciados para chequear su clonalidad.

Los oligos utilizados fueron los siguientes:

V<sub>H</sub>J558:        sentido: CGAGCTCTCCARCACAGCCTWCATGCARCTCARC.  
                  anti-sentido: GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG.

V<sub>H</sub> 7183:        sentido: CGGTACCAAGAASAMCCTGTWCCTGCAAATGASC.

anti-sentido: GTCTAGATTCTCACAAAGAGTCCGATAGACCCTGG.

V<sub>H</sub> Q52: sentido: CGGTACCAGACTGARCATCASCAAGGACAAYTCC.

anti-sentido: GTCTAGATTCTCACAAAGAGTCCGATAGACCCTGG.

DH: sentido: TTCAAAGCACAATGCCTGGCT.

anti-sentido: GTCTAGATTCTCACAAAGAGTCCGATAGACCCTGG.

C $\mu$ : sentido: TGGCCATGGGCTGCCTAGCCCGGGACTT.

anti-sentido: GCCTGACTGAGCTCACACAAGGAGGA.

### 10. PCR del transgén *Crebbp* deleciónado.

La deleción eficiente del alelo(s) condicional(es) de *Crebbp* fue confirmada mediante PCR usando los siguientes oligos:

mCBP607+324S: sentido: GGGGAAATTTTGGTCTGGTAAG

mCBP647+219AS: anti-sentido: CTGCTCTACCTAAATTCACAG

Las condiciones de la PCR consistieron en una activación inicial a 95°C (5 minutos) seguida por 36 ciclos de desnaturalización (10 segundos a 94°C), anillamiento (1 minuto a 55°C), y elongación (2 minutos a 68°C). Este paso fue seguido por una elongación final de 10 minutos a 72°C. El producto de la PCR fue corrido en un gel de agarosa al 1%, observándose los tamaños de bandas esperados para el alelo silvestre (1.1 kb), para el alelo condicional (1.3 kb) y para los alelos deleciónados (800 pb).

### 11. Análisis de clonalidad del gen del receptor de linfocitos T (TCR).

Los reordenamientos del gen del receptor de células T (*TCR*) fueron amplificados por PCR utilizando los oligos listados al final de esta sección. Las condiciones de la PCR consistieron en una activación inicial a 95°C seguida por 35 ciclos de desnaturalización (1 minuto a 95°C), anillamiento (1 minuto a 57°C), y elongación (1 minuto y 45 segundos a 72°C). Este paso fue seguido por una elongación final de 10 minutos a 72°C. Para determinar las secuencias de ADN de cada reordenamiento individual, los fragmentos de PCR fueron aislados del gel de agarosa 0,8% y clonados en el vector TA. Los insertos de ADN de, al menos, 10 clones correspondientes al mismo fragmento de la PCR, fueron secuenciados para chequear su clonalidad.

Los oligos utilizados fueron los siguientes:

D $\beta$ 2:	sentido: GTAGGCACCTGTGGGGAAGAACT anti-sentido: TGAGAGCTGTCTCCTACTATCGATT
V $\beta$ 2:	sentido: GGGTCACTGATACGGAGCTG anti-sentido: TGAGAGCTGTCTCCTACTATCGATT
V $\beta$ 4:	sentido: GGACAATCAGACTGCCTCAAGT anti-sentido: TGAGAGCTGTCTCCTACTATCGATT
V $\beta$ 5.1:	sentido: GTCCAACAGTTTGATGACTATCAC anti-sentido: TGAGAGCTGTCTCCTACTATCGATT
V $\beta$ 8:	sentido: GATGACATCATCAGGTTTTGTC anti-sentido: TGAGAGCTGTCTCCTACTATCGATT
V $\beta$ 14:	sentido: CTTCTACCTCTGTGCCTGGAGT anti-sentido: TGAGAGCTGTCTCCTACTATCGATT

### 12. Hibridación genómica comparada o CGH.

La hibridación genómica comparada se llevó a cabo en la empresa NIMGenetics de Madrid a la que se le proporcionó ADN de ratones transgénicos *Sca1-Lmo2* (12 casos de T-ALL), además de ADN de cola de un ratón control (*WT*) del mismo fondo genético. La extracción de ADN se llevó a cabo siguiendo el procedimiento descrito en el apartado “8. Extracción de ADN” pero usando timo en el caso de las leucemias T y cola en el caso del control (*WT*).

El análisis del genoma fue llevado a cabo usando el *CGH-array M.musculus whole genome 4x180k oligonucleotide array* (AMADID 27411, Agilent Technologies), siguiendo el protocolo aconsejado por el fabricante. Los datos del microarray fueron extraídos y visualizados usando el programa *Feature Extraction software* v10.7 y el *Agilent Genomic Workbench* v5.0 (Agilent Technologies). Las regiones genómicas con número de copias alterado fueron detectadas usando la estadística ADM-2 con un número mínimo de 5 sondas consecutivas. La base de datos *Genomic build mm9* fue consultada para la identificación de los posibles genes alterados.

### 13. Extracción de ARN.

Los tejidos fueron disgregados con filtros “*cell strainer*” de 70 $\mu$ m de nylon (BD FALCON). Posteriormente, se procedió a la extracción de ARN utilizando TRIzol (Life Technologies), de acuerdo con las instrucciones del fabricante.

La calidad e integridad del ARN obtenido fue calculada en función de la relación resultante del cociente entre la longitud de onda de absorción del ARN, 260nm, y de las proteínas, 280nm, utilizando para ello el espectrofotómetro Nanodrop-100 (THERMO Scientific). Los restos de fenol procedentes del uso del reactivo TRIzol fueron eliminados para evitar posibles interferencias del mismo en aplicaciones posteriores. Para ello, se usó el kit comercial “Rneasy Mini Kit” (Ref. 74104, Qiagen) y el protocolo de limpieza del ARN del mismo.

#### **14. Perfil de expresión génica**

El RNA total fue aislado en 2 pasos, utilizando TRIzol (Life Technologies) seguido de la purificación con RNeasy Mini-Kit (Qiagen) siguiendo el protocolo “RNA clean-up” con tratamiento opcional de DNase. La integridad y calidad del ARN, así como su concentración, fueron verificadas y medidas utilizando Agilent (Agilent Technologies). Las muestras fueron analizadas utilizando los *arrays* “Affymetrix Mouse Gene 1.0 y 2.0 ST”.

Brevemente, el algoritmo “*robust microarray analysis*” (RMA) fue utilizado para la corrección del *background*, la normalización intra- e inter-microarray, y el cálculo de la expresión de la señal<sup>324-326</sup>. Una vez fue calculada la señal de expresión absoluta para cada gen (por ejemplo, el valor de cada *probe set*) en cada microarray, un método denominado “*significance analysis of microarray*” (SAM)<sup>327</sup> fue utilizado para calcular la expresión diferencial significativa y encontrar los *probe sets* de genes que caracterizaban las muestras problema comparadas con las muestras control. Este método utiliza permutaciones para calcular de forma robusta la inferencia estadística de la mayoría de los genes y proporciona p-valores ajustados a múltiples test utilizando el “*false discovery rate*” (FDR)<sup>328</sup>. Un corte en, al menos, FDR<0.05 fue utilizado para los cálculos de expresión diferencial. Todos estos métodos fueron aplicados utilizando R<sup>329</sup> y Bioconductor<sup>330</sup>.

Los datos discutidos en esta tesis doctoral han sido publicados y depositados en la plataforma Gene Expression Omnibus<sup>331</sup> y son accesibles a través de los números de acceso de GEO GSE85490, GSE83570, GSE83571 y GSE83572.

#### **15. Análisis de enriquecimiento.**

El análisis de enriquecimiento de los genes expresados diferencialmente fue realizado utilizando los genes de la colección C2 de la plataforma GSEA de MSigDB<sup>332</sup> (<http://www.broad.mit.edu/gsea/>). El GSEA se realizó mediante una estadística de enriquecimiento ponderado, clasificación métrica de la señal de ruido, y corregida para

múltiples pruebas de hipótesis (MHT) utilizando 1000 permutaciones para proporcionar una tasa de falso descubrimiento corregido (FDR)  $q$ - valor.

El solapamiento de las firmas de expresión génica que son específicamente activadas o reprimidas en los *gene sets* “Módulo ESC”<sup>89</sup> y “Módulo de pluripotencia”<sup>333</sup> fue evaluado con los genes activados y reprimidos de las muestras problema utilizando el análisis de enriquecimiento de GSEA. Del mismo modo fue evaluado el solapamiento del *gene set* de los genes de la vía de señalización de *Notch1* (obtenidos de <http://www.broad.mit.edu/gsea/>).

### 16. Secuenciación mediante inmunoprecipitación de la cromatina (ChIP-seq)

Se llevó a cabo ChIP-seq de 100.000 células B purificadas de ratones *Cbp*<sup>F/F</sup> x *E $\mu$ Bcl2* y de ratones *Cbp* <sup>$\Delta\Delta$</sup>  x *E $\mu$ Bcl2*. Estas células fueron fijadas en formaldehído 1%, se añadió glicina para parar el entrecruzamiento, y finalmente las células fueron lavadas y lisadas. La cromatina fue cortada por sonicación usando el aparato Covaris S2, obteniendo fragmentos de entre 100-300 pb, que fueron confirmados mediante Agilent TapeStation. La inmunoprecipitación de la cromatina se llevó a cabo usando el anticuerpo anti-H3K18Ac (Cell Signaling Technologies, Ref. 9675) y el instrumento IPstar de Diagenode. El ADN fue purificado usando columnas MicroChIP DiaPure de Diagenode (Ref. C03040001). Las librerías TruSeq fueron construidas usando el kit de preparación de librerías MicroPlex<sup>TM</sup> de Diagenode (Ref. C05010012) y el aparato IPstart siguiendo el protocolo facilitado por el fabricante. Cada librería fue amplificada a través de 10 ciclos de PCR y cada lote de 4 muestras fue reunido en cantidades equimolares para su secuenciación. Cada grupo fue secuenciado con lecturas finales emparejadas de 100 pb en un modo de alto rendimiento usando el aparato Illumina 2500.

Las secuencias en bruto iniciales fueron procesadas según los siguientes pasos. Las secuencias adaptadoras se recortaron y los archivos *fastq* resultantes se alinearon con el genoma de ratón (NCBI37/mm9) utilizando el alineador de secuencias Bowtie2 (versión 2.2.3)<sup>334</sup>. Los duplicados se eliminaron utilizando las herramientas SAM MarkDuplicates<sup>335</sup> y los archivos *bam* fueron ordenados y convertidos al formato *sam* para la búsqueda de picos. Se compararon las muestras del mismo lote utilizando el software “ChIP-seq pipeline Fish the ChIPs”<sup>336</sup> que emplea el robusto algoritmo de búsqueda de picos ChIP-Seq MACS<sup>337</sup>. Como las réplicas de ChIPseq se realizaron en lotes, con una muestra de *Cbp* <sup>$\Delta\Delta$</sup>  x *E $\mu$ Bcl2* y una de *Cbp*<sup>F/F</sup> x *E $\mu$ Bcl2* y sus controles de entrada coincidentes en cada lote, se realizaron comparaciones uno a

uno de las muestras dentro de cada lote para identificar regiones con un incremento o reducción significativa (FDR <0.25) de los niveles de acetilación. Después se derivó la intersección entre cada uno de los dos lotes utilizando el software EaSeq<sup>338</sup>. Los datos públicos de ChIP-seq disponibles para Myc de la línea celular de linfoma de células B murinas, Ch12, y la línea celular de eritroleucemia murina, MEL, se descargaron en formato de archivo *bam*, mapeado a mm9, desde la base de datos ENCODE<sup>339</sup>. Los mapas de calor de ChIP-seq se construyeron utilizando archivos *bam* deduplicados con el software EaSeq<sup>338</sup>.

### **17. Análisis de enriquecimiento de los motivos de ADN.**

Las secuencias de nucleótidos de las regiones con una alteración significativa de H3K18Ac en las células B de los ratones *Cbp<sup>ΔΔ</sup> x EμBcl2* en comparación con las células B de los ratones *Cbp<sup>F/F</sup> x EμBcl2* fueron descargadas usando el navegador del genoma de la Universidad de California<sup>340</sup>. Como control, utilizamos regiones de tamaño similar que se localizaban inmediatamente aguas arriba e inmediatamente agua abajo. Las secuencias de nucleótidos fueron analizadas usando CisFinder<sup>341</sup> para identificar motivos significativamente sobrerrepresentados, usando la configuración predeterminada (FDR>0.1). Aquellos motivos con un 75% de similitud fueron agrupados y cada grupo fue comparado con matrices de secuencias de ADN para 348 factores de transcripción conocidos en la base de datos CisView utilizando el parámetro predeterminado del 75% de similitud<sup>342</sup>.

### **18. Análisis de hibridación fluorescente *in situ* (FISH).**

La hibridación fluorescente *in situ* (FISH) de los pacientes con T-ALL *LMO2+* se llevó a cabo usando dos clones BAC que flanqueaban el extremo 5' (CY3 RP11-98C11 en rojo) y el extremo 3' (FITC RP11-60G13 en verde) del gen *LMO2*, siguiendo el procedimiento descrito previamente<sup>343</sup>.

El estado de la translocación de *MYC* se determinó mediante FISH en un subconjunto de casos de DLBCL humano, como parte de la rutina de atención clínica en el Centro Médico de la Universidad de Nebraska (UNMC), siguiendo el protocolo descrito previamente<sup>344</sup>.

### 19. Preparación de librerías del exoma completo humano y de ratón (WES).

#### 19.1 Secuenciación de nueva generación (NGS) de los linfomas de células B.

Para los tumores murinos, se extrajo el ADN de bazos implicados en linfoma y de bazos de ratones control sanos de la misma edad, siguiendo el procedimiento explicado en detalle en la sección “8. Extracción de ADN”.

Todos los estudios de muestras de pacientes humanos fueron aprobados por las respectivas juntas institucionales de revisión. Así, se obtuvieron biopsias rápidas pretratamiento congeladas de 126 tumores de FL y 140 de DLBCL (cohorte del Centro Médico de la Universidad de Nebraska, UNMC). Para los tumores humanos, el tejido se homogeneizó y el ADN y el ARN se extrajeron utilizando el Kit AllPrep Mini de Qiagen (Ref. 80204) de acuerdo con el protocolo del fabricante.

Para todos los tumores, se cortó entre 1 y 2 µg de ADN mediante sonicación con un instrumento Covaris S2, y las bibliotecas de secuenciación de nueva generación fueron preparadas usando los kits KAPA Hyper Prep (de KAPA Biosystems) y los adaptadores Illumina TruSeq. Las bibliotecas se amplificaron un máximo de 8 ciclos de PCR antes de la agrupación equimolar. Las bibliotecas de tumores humanos se agruparon en lotes de 12, y las bibliotecas de tumores murinos en lotes de 4.

Para los tumores humanos, se realizó la captura híbrida de 380 genes a una profundidad media de 634X (mínimo 263X, máximo 1396X) usando reactivos SeqCap de Nimblegen como se describió previamente<sup>143</sup>. Los tumores murinos fueron enriquecidos para el exoma codificante utilizando los reactivos “*SeqCap Custom Mouse Exome Capture*” (de Nimblegen).

Después de la captura, las agrupaciones se amplificaron con 10 ciclos de PCR y se secuenciaron con lecturas finales emparejadas de 100 pb en un modo de alto rendimiento en el aparato Illumina HiSeq 2500 en el Instituto Biotecnológico Hudson Alpha.

Los datos se analizaron utilizando nuestra línea de análisis de secuenciación interna. En resumen, se evaluó la calidad de los archivos FASTQ utilizando FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) y los datos se alinearon con el genoma de referencia UCG hg19 o mm10 utilizando BWA-Mem<sup>345</sup>. Las lecturas mapeadas se realinearon alrededor de InDels utilizando el realineador GATK InDel<sup>346</sup>,

ordenadas y deduplicadas usando las herramientas SAM<sup>335</sup>. Las variantes se determinaron utilizando VarScan2<sup>347</sup> y GATK Unified Genotyper<sup>346</sup>, y los exomas humanos se anotaron usando SeattleSeq<sup>348</sup>. Los exomas de ratón se anotaron usando Annovar<sup>349</sup>. El filtrado de variantes de línea germinal y de variantes en regiones de repeticiones en tándem se realizó como se describió previamente<sup>143</sup>, y sólo se consideraron variantes aquellas determinadas por ambos algoritmos y respaldadas por  $\geq 5$  lecturas en cada dirección. Por lo tanto, se incluyeron los eventos clonales y subclonales.

Los datos de secuenciación de los 138 FL previamente secuenciados<sup>143</sup> también fueron reanalizados siguiendo la misma línea de análisis para asegurar la consistencia en los resultados obtenidos. Los datos de mutaciones en *CREBBP* de 46 FL y de 134 DLBCL fueron obtenidos de un estudio previo<sup>350</sup>. Los datos del perfil de expresión génica de los 73 DLBCL se obtuvieron de la plataforma Gene Expression Omnibus<sup>331</sup> (GSE12195).

### **19.2 Secuenciación de nueva generación (NGS) de las T-ALL humanas y murinas.**

El ADN de las células T leucémicas fue purificado siguiendo los métodos detallados en la sección “8. Extracción de ADN”.

La preparación de la biblioteca del exoma fue realizada utilizando el kit Agilent *SureSelectXT Mouse All exon*. Se realizó la captura dirigida por la hibridación con una biblioteca de ARN, siguiendo el protocolo proporcionado por el fabricante. La purificación y el enriquecimiento de la biblioteca capturada fue conseguida mediante la unión a “*MyOne Streptavidin T1 Dynabeads*” (Life Technologies) y la amplificación de los elementos no unidos a los *beads* por PCR en un rango lineal. La secuenciación de  $2 \times 10^2$  pb con un índice de lectura de 6 pb fue realizada utilizando el kit TruSeq SBS kit v3 en el secuenciador HiSeq 2500 (Illumina).

Los archivos fastq fueron generados utilizando BcltoFastq 1.8.4 (Illumina). Se utilizó la versión 0.7.4. de BWA para alinear los datos secuenciados con el genoma de referencia de ratón (GRCm38.71). Los pasos de conversión fueron llevados a cabo utilizando Samtools, seguidos por la retirada de las lecturas duplicadas (<http://broadinstitute.github.io/picard>). El reordenamiento local alrededor de los indels, la denominación de los SNP, la anotación y la recalibración fueron facilitados por GATK 2.4.9. Los SNPs de ratón dbSNP138 y dbSNP fueron utilizados como conjuntos de datos de prueba para las cepas de ratón utilizadas. Las variaciones resultantes

entre los datos secuenciados y la secuencia referencia fueron anotadas por el programa Variant Effect Predictor utilizando la base de datos de Ensembl (v70) e importadas a una base de datos propia (servidor MySQL interno), para facilitar la anotación automática y manual, y el análisis de los datos mediante consultas complejas de bases de datos.

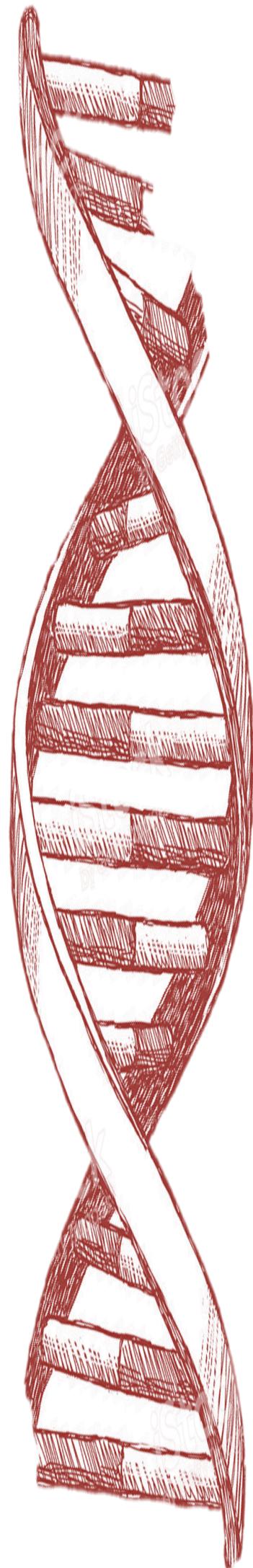
Las anotaciones somáticas fueron realizadas utilizando MuTect<sup>351</sup>. Los genes relacionados con cáncer fueron determinados traduciendo la lista “*cancer gene consensus*” de COSMIC, utilizando la aplicación Biomart de Ensembl. Para evaluar deleciones en las muestras tumorales, las secuencias brutas fueron mapeadas con el genoma de referencia mm10/GRCm38 con BWA-MEM<sup>345</sup> (parámetros por defecto, versión 0.7.10) y las lecturas con extremos finales duplicados fueron eliminadas basándonos en sus posiciones de mapeo. Los alineamientos fueron analizados con SoftSV<sup>352</sup> (parámetros por defecto, versión 1.4) para anotar variaciones estructurales (pequeñas deleciones, inversiones, duplicaciones en tándem y translocaciones) y con Control-FREEC<sup>353</sup> (versión 7.2) para detectar alteraciones en el *copy number*.

Las muestras humanas de médula ósea con translocación t(11;14)(q13;q11) fueron adquiridas después del consentimiento informado en el estudio AIEOP-BFM-ALL y procesado a través de la plataforma Macrogen Europe, Heidelberg.





*Anexos*





ANEXOS

Tejido	Subgrupo celular	Cbp WT/F			EμBcl2			Cbp WT/A			Cbp A/A			Cbp WT/A x EμBcl2			Cbp A/A x EμBcl2		
		Media	SD	N	Media	SD	N	Media	SD	N	Media	SD	N	Media	SD	N	Media	SD	N
Médula ósea	Células B totales	31.35	8.64	4	21.00	10.12	3	22.55	8.64	4	14.72	6.39	3	76.17	12.87	4	44.34	2.65	4
Médula ósea	Células proB y preB	12.90	2.81	4	3.33	1.43	3	10.14	6.78	4	7.64	5.19	3	17.53	5.18	4	13.46	6.41	4
Médula ósea	Células B inmaduras	2.59	0.76	4	1.11	1.21	3	1.60	1.20	4	0.46	0.32	3	15.35	5.36	4	6.12	1.52	4
Médula ósea	Células B maduras	2.93	2.37	4	3.85	2.12	3	2.75	0.45	4	1.40	0.36	3	18.33	5.33	4	7.42	2.89	4
Médula ósea	Células B recirculantes	8.10	1.61	4	8.13	3.36	3	4.99	1.19	4	3.60	0.87	3	10.90	2.89	4	12.61	4.78	4
Médula ósea	Células T CD8	1.49	0.67	4	0.89	0.54	3	1.04	0.28	4	1.14	0.29	3	1.11	0.13	4	2.42	0.25	4
Médula ósea	Células T CD4	1.46	0.39	4	2.30	1.17	3	1.33	0.48	4	1.25	0.21	3	1.77	0.35	4	3.49	0.82	4
Médula ósea	Células mieloides	35.03	8.86	4	63.33	9.14	3	48.05	22.69	4	51.87	14.60	3	20.50	6.35	4	23.85	3.88	4
Bazo	Células B totales	45.83	9.18	4	38.43	10.20	3	29.95	4.52	4	18.84	5.14	3	60.73	2.43	4	31.85	8.17	4
Bazo	Células B B220+IgM-	8.15	1.89	4	13.49	4.41	3	7.79	0.88	4	5.78	0.99	3	24.03	4.64	4	13.92	4.43	4
Bazo	Células B B220+IgM+	37.85	7.68	4	25.53	6.43	3	22.33	4.64	4	13.07	4.27	3	36.70	3.27	4	17.92	4.69	4
Bazo	Células B foliculares	33.74	4.69	4	23.19	8.11	3	18.46	2.97	4	9.88	1.70	3	41.03	2.13	4	16.30	5.08	4
Bazo	Células B de la zona marginal	4.20	0.74	4	1.90	0.29	3	2.32	0.84	4	2.29	0.91	3	2.19	0.48	4	2.28	0.69	4
Bazo	Células T CD8	21.45	4.48	4	15.07	6.69	3	27.38	4.96	4	18.50	2.03	3	10.51	1.04	4	22.23	5.64	4
Bazo	Células T CD4	14.65	4.71	4	16.57	3.84	3	17.40	2.17	4	28.53	5.08	3	13.30	1.68	4	22.73	3.38	4
Bazo	Células mieloides	2.65	1.46	4	2.73	0.40	3	5.06	7.17	4	12.84	9.92	3	0.59	0.19	4	0.99	0.22	4

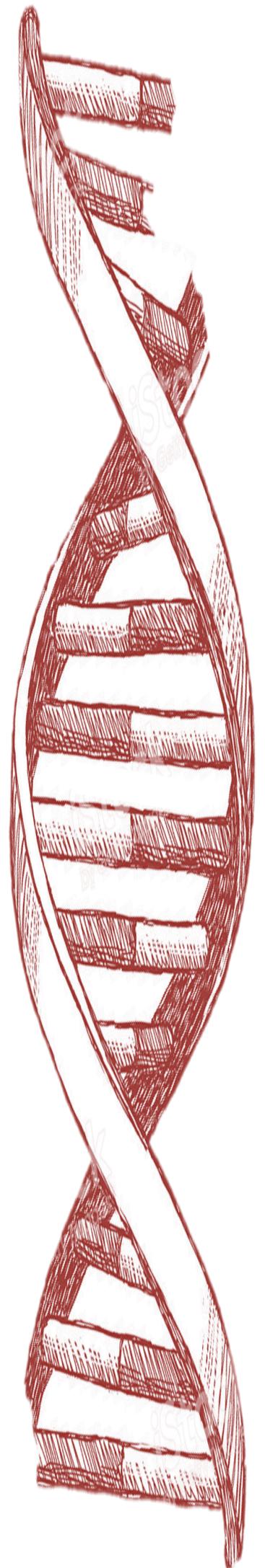
**Anexo 1: Tabla resumen de los estudios de citometría de flujo realizados en los ratones jóvenes.** Para estudiar el papel de Crebbp en el desarrollo de células B a corto plazo, se llevaron a cabo análisis de citometría de flujo de los diferentes compartimentos celulares hematopoyéticos de ratones de entre 11-20 semanas de edad de todas las cepas en estudio. Para cada cepa en estudio se muestra el porcentaje medio de cada subtipo celular de médula ósea y bazo, la desviación estándar (SD) y el número de ratones incluidos (N).

Tejido	Subtipo celular	Cbp <sup>WT/F</sup>			EuBcl2			Cbp <sup>WT/A</sup>			Cbp <sup>d/A</sup>			Cbp <sup>WT/A</sup> x EuBcl2			Cbp <sup>d/A</sup> x EuBcl2		
		Media	SD	N	Media	SD	N	Media	SD	N	Media	SD	N	Media	SD	N	Media	SD	N
Médula ósea	B220+	19.78	1.56	10	26.88	1.68	11	12.86	1.28	33	15.04	3.20	20	17.35	1.95	24	24.49	3.03	14
Médula ósea	Células B totales	8.26	1.33	10	6.58	0.84	11	5.29	0.86	33	5.32	1.12	20	5.76	0.80	24	5.81	1.21	14
Médula ósea	Células proB y preB	2.12	0.33	10	0.43	0.20	11	1.10	0.19	33	3.81	3.20	20	3.00	0.66	24	8.83	3.29	14
Médula ósea	Células B inmaduras	4.81	0.52	10	12.03	1.08	11	3.76	0.44	33	3.42	1.39	20	5.07	0.74	24	5.43	1.30	14
Médula ósea	Células B maduras	9.40	0.82	10	19.87	1.20	11	6.47	0.70	33	5.90	1.63	20	8.58	0.85	24	9.86	1.30	14
Médula ósea	Células B recirculantes	1.35	0.20	10	3.73	0.62	11	2.11	0.30	33	1.55	0.26	20	2.68	0.28	24	2.77	0.27	14
Médula ósea	Células T CD8	1.44	0.10	10	3.21	0.31	11	1.59	0.13	33	1.12	0.08	20	3.10	0.28	24	4.78	1.25	14
Médula ósea	Células T CD4	44.35	2.07	10	30.01	2.14	11	49.25	2.04	33	50.97	3.03	20	37.88	1.78	24	28.68	2.91	14
Médula ósea	Células mieloides	52.22	2.81	10	49.21	2.07	11	34.56	3.08	33	32.21	5.60	20	32.91	2.76	24	35.58	5.67	14
Bazo	B220+	5.05	0.43	10	4.07	0.22	11	1.63	0.31	33	2.25	0.43	20	1.86	0.33	24	5.79	4.03	14
Bazo	B220+ IgM <sup>fuerte</sup> IgD-	11.25	1.02	10	11.94	0.68	11	4.79	0.87	33	4.84	0.90	20	5.41	0.80	24	3.73	0.74	14
Bazo	B220+ IgM <sup>fuerte</sup> IgD+	19.49	1.72	10	19.53	1.27	11	16.97	1.75	33	11.81	3.22	20	10.09	1.19	24	5.20	1.16	14
Bazo	B220+ IgM+ IgD+	34.72	2.63	10	32.11	1.81	11	18.12	2.11	33	11.92	2.40	20	15.82	2.07	24	9.28	1.87	14
Bazo	B220+ CD21+ CD23+	3.20	0.20	10	4.25	0.50	11	1.85	0.31	33	2.78	0.57	20	2.61	0.44	24	2.32	0.58	14
Bazo	Células B de la zona marginal	37.05	2.51	10	38.77	2.35	11	24.31	2.50	33	18.94	3.82	20	18.85	2.30	24	14.14	4.10	14
Bazo	Células B recirculantes	11.65	1.13	10	16.25	1.61	11	10.54	1.11	33	15.04	1.92	20	17.02	1.66	24	17.05	2.78	14
Bazo	Células T CD8	15.91	1.55	10	12.95	0.78	11	14.09	1.25	33	17.34	2.11	20	15.54	1.24	24	20.09	3.52	14
Bazo	Células T CD4	2.95	0.63	10	3.25	0.62	11	6.01	1.42	33	6.20	1.62	20	4.48	1.27	24	2.20	0.42	14
Bazo	Células mieloides																		

**Anexo 1: Tabla resumen de los estudios de citometría de flujo realizados en los ratones adultos.** Para estudiar el papel de *Crebpb* en el desarrollo de células B a largo plazo, se llevaron a cabo análisis de citometría de flujo de los diferentes compartimentos celulares hematopoyéticos de ratones de más de 5 meses de edad de todas las cepas en estudio. Para cada cepa en estudio se muestra el porcentaje medio de cada subtipo celular de médula ósea y bazo, la desviación estándar (SD) y el número de ratones incluidos (N). Además se incluye una columna con el inmunofenotipo específico de cada subtipo celular.







# *Bibliografía*



## BIBLIOGRAFÍA

- 1 Ferlay J, S. I., Ervik M, Dikshit R, Eser S, Mathers C et al. . in *IARC CancerBase* Vol. 11 (2012).
- 2 Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
- 3 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
- 4 Sanchez-Garcia, I., Vicente-Duenas, C. & Cobaleda, C. The theoretical basis of cancer-stem-cell-based therapeutics of cancer: can it be put into practice? *BioEssays : news and reviews in molecular, cellular and developmental biology* **29**, 1269-1280, doi:10.1002/bies.20679 (2007).
- 5 Etzioni, R. *et al.* The case for early detection. *Nature reviews. Cancer* **3**, 243-252, doi:10.1038/nrc1041 (2003).
- 6 Jemal, A. *et al.* Cancer statistics, 2008. *CA: a cancer journal for clinicians* **58**, 71-96, doi:10.3322/CA.2007.0010 (2008).
- 7 Vicente-Duenas, C., Romero-Camarero, I., Cobaleda, C. & Sanchez-Garcia, I. Function of oncogenes in cancer development: a changing paradigm. *The EMBO journal* **32**, 1502-1513, doi:10.1038/emboj.2013.97 (2013).
- 8 Der, C. J., Krontiris, T. G. & Cooper, G. M. Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 3637-3640 (1982).
- 9 Goldfarb, M., Shimizu, K., Perucho, M. & Wigler, M. Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. *Nature* **296**, 404-409 (1982).
- 10 Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene. *Nature* **297**, 474-478 (1982).
- 11 Pulciani, S. *et al.* Oncogenes in human tumor cell lines: molecular cloning of a transforming gene from human bladder carcinoma cells. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 2845-2849 (1982).
- 12 Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S. & Barbacid, M. T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes. *Nature* **298**, 343-347 (1982).
- 13 Shih, C. & Weinberg, R. A. Isolation of a transforming sequence from a human bladder carcinoma cell line. *Cell* **29**, 161-169 (1982).
- 14 Chin, L. *et al.* Essential role for oncogenic Ras in tumour maintenance. *Nature* **400**, 468-472, doi:10.1038/22788 (1999).

- 15 Huettner, C. S., Zhang, P., Van Etten, R. A. & Tenen, D. G. Reversibility of acute B-cell leukaemia induced by BCR-ABL1. *Nature genetics* **24**, 57-60, doi:10.1038/71691 (2000).
- 16 Boxer, R. B., Jang, J. W., Sintasath, L. & Chodosh, L. A. Lack of sustained regression of c-MYC-induced mammary adenocarcinomas following brief or prolonged MYC inactivation. *Cancer cell* **6**, 577-586, doi:10.1016/j.ccr.2004.10.013 (2004).
- 17 Weinstein, I. B. Cancer. Addiction to oncogenes--the Achilles heal of cancer. *Science* **297**, 63-64, doi:10.1126/science.1073096 (2002).
- 18 Perez-Caro, M. *et al.* Cancer induction by restriction of oncogene expression to the stem cell compartment. *The EMBO journal* **28**, 8-20, doi:10.1038/emboj.2008.253 (2009).
- 19 Vicente-Duenas, C. *et al.* Expression of MALT1 oncogene in hematopoietic stem/progenitor cells recapitulates the pathogenesis of human lymphoma in mice. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 10534-10539, doi:10.1073/pnas.1204127109 (2012).
- 20 Vicente-Duenas, C. *et al.* A novel molecular mechanism involved in multiple myeloma development revealed by targeting MafB to haematopoietic progenitors. *The EMBO journal* **31**, 3704-3717, doi:10.1038/emboj.2012.227 (2012).
- 21 Romero-Camarero, I. *et al.* Germinal centre protein HGAL promotes lymphoid hyperplasia and amyloidosis via BCR-mediated Syk activation. *Nature communications* **4**, 1338, doi:10.1038/ncomms2334 (2013).
- 22 Rodriguez-Hernandez, G. *et al.* Infection Exposure Promotes ETV6-RUNX1 Precursor B-cell Leukemia via Impaired H3K4 Demethylases. *Cancer research* **77**, 4365-4377, doi:10.1158/0008-5472.CAN-17-0701 (2017).
- 23 Green, M. R. *et al.* Transient expression of Bcl6 is sufficient for oncogenic function and induction of mature B-cell lymphoma. *Nature communications* **5**, 3904, doi:10.1038/ncomms4904 (2014).
- 24 F, M., B, J. & F, M. (National Cancer Institute, 2015).
- 25 Fialkow, P. J., Jacobson, R. J. & Papayannopoulou, T. Chronic myelocytic leukemia: clonal origin in a stem cell common to the granulocyte, erythrocyte, platelet and monocyte/macrophage. *The American journal of medicine* **63**, 125-130 (1977).
- 26 Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105-111, doi:10.1038/35102167 (2001).
- 27 Dalerba, P., Cho, R. W. & Clarke, M. F. Cancer stem cells: models and concepts. *Annual review of medicine* **58**, 267-284, doi:10.1146/annurev.med.58.062105.204854 (2007).
- 28 Shackleton, M., Quintana, E., Fearon, E. R. & Morrison, S. J. Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* **138**, 822-829, doi:10.1016/j.cell.2009.08.017 (2009).

- 29 Zellmer, V. R. & Zhang, S. Evolving concepts of tumor heterogeneity. *Cell & bioscience* **4**, 69, doi:10.1186/2045-3701-4-69 (2014).
- 30 Furth, J. The transmission of leukemia of mice with a single cell. *The American Journal of Cancer* **31**, 276–282 (1937).
- 31 Park, C. H., Bergsagel, D. E. & McCulloch, E. A. Mouse myeloma tumor stem cells: a primary cell culture assay. *Journal of the National Cancer Institute* **46**, 411-422 (1971).
- 32 Hamburger, A. & Salmon, S. E. Primary bioassay of human myeloma stem cells. *The Journal of clinical investigation* **60**, 846-854, doi:10.1172/JCI108839 (1977).
- 33 Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature medicine* **3**, 730-737 (1997).
- 34 Miyamoto, T., Weissman, I. L. & Akashi, K. AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7521-7526 (2000).
- 35 Cobaleda, C. *et al.* A primitive hematopoietic cell is the target for the leukemic transformation in human philadelphia-positive acute lymphoblastic leukemia. *Blood* **95**, 1007-1013 (2000).
- 36 Cox, C. V. *et al.* Characterization of acute lymphoblastic leukemia progenitor cells. *Blood* **104**, 2919-2925, doi:10.1182/blood-2004-03-0901 (2004).
- 37 Cox, C. V. *et al.* Characterization of a progenitor cell population in childhood T-cell acute lymphoblastic leukemia. *Blood* **109**, 674-682, doi:10.1182/blood-2006-06-030445 (2007).
- 38 Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3983-3988, doi:10.1073/pnas.0530291100 (2003).
- 39 Singh, S. K. *et al.* Identification of a cancer stem cell in human brain tumors. *Cancer research* **63**, 5821-5828 (2003).
- 40 Li, C. *et al.* Identification of pancreatic cancer stem cells. *Cancer research* **67**, 1030-1037, doi:10.1158/0008-5472.CAN-06-2030 (2007).
- 41 Takaishi, S. *et al.* Identification of gastric cancer stem cells using the cell surface marker CD44. *Stem cells* **27**, 1006-1020, doi:10.1002/stem.30 (2009).
- 42 Yamashita, T. & Wang, X. W. Cancer stem cells in the development of liver cancer. *The Journal of clinical investigation* **123**, 1911-1918, doi:10.1172/JCI66024 (2013).
- 43 Han, J., Fujisawa, T., Husain, S. R. & Puri, R. K. Identification and characterization of cancer stem cells in human head and neck squamous cell carcinoma. *BMC cancer* **14**, 173, doi:10.1186/1471-2407-14-173 (2014).

- 44 Cammareri, P. *et al.* Isolation and culture of colon cancer stem cells. *Methods in cell biology* **86**, 311-324, doi:10.1016/S0091-679X(08)00014-9 (2008).
- 45 Hurt, E. M., Kawasaki, B. T., Klarmann, G. J., Thomas, S. B. & Farrar, W. L. CD44+ CD24(-) prostate cells are early cancer progenitor/stem cells that provide a model for patients with poor prognosis. *British journal of cancer* **98**, 756-765, doi:10.1038/sj.bjc.6604242 (2008).
- 46 Nassar, D. & Blanpain, C. Cancer Stem Cells: Basic Concepts and Therapeutic Implications. *Annual review of pathology* **11**, 47-76, doi:10.1146/annurev-pathol-012615-044438 (2016).
- 47 Beck, B. & Blanpain, C. Unravelling cancer stem cell potential. *Nature reviews. Cancer* **13**, 727-738, doi:10.1038/nrc3597 (2013).
- 48 Li, L., Borodyansky, L. & Yang, Y. Genomic instability en route to and from cancer stem cells. *Cell cycle* **8**, 1000-1002, doi:10.4161/cc.8.7.8041 (2009).
- 49 Rapp, U. R., Ceteci, F. & Schreck, R. Oncogene-induced plasticity and cancer stem cells. *Cell cycle* **7**, 45-51, doi:10.4161/cc.7.1.5203 (2008).
- 50 Romero-Camarero, I. *et al.* New models towards assessing anti-cancer therapeutics. *Histology and histopathology* **27**, 157-170, doi:10.14670/HH-27.157 (2012).
- 51 Polakis, P. The oncogenic activation of beta-catenin. *Current opinion in genetics & development* **9**, 15-21 (1999).
- 52 Willert, K. *et al.* Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **423**, 448-452, doi:10.1038/nature01611 (2003).
- 53 Reya, T. *et al.* A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* **423**, 409-414, doi:10.1038/nature01593 (2003).
- 54 Pear, W. S. *et al.* Exclusive development of T cell neoplasms in mice transplanted with bone marrow expressing activated Notch alleles. *The Journal of experimental medicine* **183**, 2283-2291 (1996).
- 55 Varnum-Finney, B. *et al.* Pluripotent, cytokine-dependent, hematopoietic stem cells are immortalized by constitutive Notch1 signaling. *Nature medicine* **6**, 1278-1281, doi:10.1038/81390 (2000).
- 56 Wetmore, C. Sonic hedgehog in normal and neoplastic proliferation: insight gained from human tumors and animal models. *Current opinion in genetics & development* **13**, 34-42 (2003).
- 57 Park, I. K. *et al.* Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423**, 302-305, doi:10.1038/nature01587 (2003).
- 58 Lessard, J. & Sauvageau, G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. *Nature* **423**, 255-260, doi:10.1038/nature01572 (2003).
- 59 Di Cristofano, A. & Pandolfi, P. P. The multiple roles of PTEN in tumor suppression. *Cell* **100**, 387-390 (2000).

- 60 Groszer, M. *et al.* Negative regulation of neural stem/progenitor cell proliferation by the Pten tumor suppressor gene in vivo. *Science* **294**, 2186-2189, doi:10.1126/science.1065518 (2001).
- 61 Pardal, R., Clarke, M. F. & Morrison, S. J. Applying the principles of stem-cell biology to cancer. *Nature reviews. Cancer* **3**, 895-902, doi:10.1038/nrc1232 (2003).
- 62 Krivtsov, A. V. *et al.* Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442**, 818-822, doi:10.1038/nature04980 (2006).
- 63 Yilmaz, O. H. *et al.* Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature* **441**, 475-482, doi:10.1038/nature04703 (2006).
- 64 Shipitsin, M. *et al.* Molecular definition of breast tumor heterogeneity. *Cancer cell* **11**, 259-273, doi:10.1016/j.ccr.2007.01.013 (2007).
- 65 Flores, M. *et al.* Recurrent DNA inversion rearrangements in the human genome. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 6099-6106, doi:10.1073/pnas.0701631104 (2007).
- 66 I, S.-G. Genetics of Cancer: Away from the beaten track. *Current Genomics* **3**, doi:10.2174/1389202023350390 (2002).
- 67 Perez-Caro, M., Perez-Mancera, P., Voces, F. & Sanchez-Garcia, I. Of Man in Mouse: Modelling Human Cancer Genotype-Phenotype Correlations in Mice. *Current Genomics* **6**, 8, doi:10.2174/1389202053642276 (2005).
- 68 Zambrowicz, B. P. & Sands, A. T. Knockouts model the 100 best-selling drugs--will they model the next 100? *Nature reviews. Drug discovery* **2**, 38-51, doi:10.1038/nrd987 (2003).
- 69 Rangarajan, A. & Weinberg, R. A. Opinion: Comparative biology of mouse versus human cells: modelling human cancer in mice. *Nature reviews. Cancer* **3**, 952-959, doi:10.1038/nrc1235 (2003).
- 70 Frese, K. K. & Tuveson, D. A. Maximizing mouse cancer models. *Nature reviews. Cancer* **7**, 645-658, doi:10.1038/nrc2192 (2007).
- 71 Walrath, J. C., Hawes, J. J., Van Dyke, T. & Reilly, K. M. Genetically engineered mouse models in cancer research. *Advances in cancer research* **106**, 113-164, doi:10.1016/S0065-230X(10)06004-5 (2010).
- 72 Cheon, D. J. & Orsulic, S. Mouse models of cancer. *Annual review of pathology* **6**, 95-119, doi:10.1146/annurev.pathol.3.121806.154244 (2011).
- 73 Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645-648, doi:10.1038/367645a0 (1994).
- 74 Visvader, J. E. Cells of origin in cancer. *Nature* **469**, 314-322, doi:10.1038/nature09781 (2011).

- 75 Vicente-Duenas, C., Perez-Caro, M., Abollo-Jimenez, F., Cobaleda, C. & Sanchez-Garcia, I. Stem-cell driven cancer: "hands-off" regulation of cancer development. *Cell cycle* **8**, 1314-1318, doi:10.4161/cc.8.9.8217 (2009).
- 76 Barker, N. The canonical Wnt/beta-catenin signalling pathway. *Methods in molecular biology* **468**, 5-15, doi:10.1007/978-1-59745-249-6\_1 (2008).
- 77 Barker, N. *et al.* Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608-611, doi:10.1038/nature07602 (2009).
- 78 Zhu, L. *et al.* Prominin 1 marks intestinal stem cells that are susceptible to neoplastic transformation. *Nature* **457**, 603-607, doi:10.1038/nature07589 (2009).
- 79 Alcantara Llaguno, S. *et al.* Malignant astrocytomas originate from neural stem/progenitor cells in a somatic tumor suppressor mouse model. *Cancer cell* **15**, 45-56, doi:10.1016/j.ccr.2008.12.006 (2009).
- 80 Dirks, P. B. Cancer's source in the peripheral nervous system. *Nature medicine* **14**, 373-375, doi:10.1038/nm0408-373 (2008).
- 81 Zheng, H. *et al.* Pten and p53 converge on c-Myc to control differentiation, self-renewal, and transformation of normal and neoplastic stem cells in glioblastoma. *Cold Spring Harb Symp Quant Biol* **73**, 427-437, doi:10.1101/sqb.2008.73.047 (2008).
- 82 Vicente-Duenas, C., Gutierrez de Diego, J., Rodriguez, F. D., Jimenez, R. & Cobaleda, C. The role of cellular plasticity in cancer development. *Current medicinal chemistry* **16**, 3676-3685 (2009).
- 83 Huntly, B. J. *et al.* MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. *Cancer cell* **6**, 587-596, doi:10.1016/j.ccr.2004.10.015 (2004).
- 84 Somervaille, T. C. & Cleary, M. L. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer cell* **10**, 257-268, doi:10.1016/j.ccr.2006.08.020 (2006).
- 85 Cozzio, A. *et al.* Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. *Genes & development* **17**, 3029-3035, doi:10.1101/gad.1143403 (2003).
- 86 So, C. W. *et al.* MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. *Cancer cell* **3**, 161-171 (2003).
- 87 Guibal, F. C. *et al.* Identification of a myeloid committed progenitor as the cancer-initiating cell in acute promyelocytic leukemia. *Blood* **114**, 5415-5425, doi:10.1182/blood-2008-10-182071 (2009).
- 88 Wojiski, S. *et al.* PML-RARalpha initiates leukemia by conferring properties of self-renewal to committed promyelocytic progenitors. *Leukemia* **23**, 1462-1471, doi:10.1038/leu.2009.63 (2009).
- 89 Wong, D. J. *et al.* Module map of stem cell genes guides creation of epithelial cancer stem cells. *Cell stem cell* **2**, 333-344, doi:10.1016/j.stem.2008.02.009 (2008).

- 90 Chen, L. *et al.* Precancerous stem cells have the potential for both benign and malignant differentiation. *PloS one* **2**, e293, doi:10.1371/journal.pone.0000293 (2007).
- 91 White, A. C. & Lowry, W. E. Refining the role for adult stem cells as cancer cells of origin. *Trends in cell biology* **25**, 11-20, doi:10.1016/j.tcb.2014.08.008 (2015).
- 92 Garcia-Ramirez, I. *et al.* Crebbp loss cooperates with Bcl2 over-expression to promote lymphoma in mice. *Blood*, doi:10.1182/blood-2016-08-733469 (2017).
- 93 Castellanos, A. *et al.* Cancer as a reprogramming-like disease: implications in tumor development and treatment. *Seminars in cancer biology* **20**, 93-97, doi:10.1016/j.semcan.2010.02.001 (2010).
- 94 Hanna, J., Carey, B. W. & Jaenisch, R. Reprogramming of somatic cell identity. *Cold Spring Harb Symp Quant Biol* **73**, 147-155, doi:10.1101/sqb.2008.73.025 (2008).
- 95 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).
- 96 Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872, doi:10.1016/j.cell.2007.11.019 (2007).
- 97 Cobaleda, C. & Sanchez-Garcia, I. Back to the beginning: The initiation of cancer. *BioEssays : news and reviews in molecular, cellular and developmental biology* **35**, 413, doi:10.1002/bies.201300024 (2013).
- 98 Krizhanovsky, V. & Lowe, S. W. Stem cells: The promises and perils of p53. *Nature* **460**, 1085-1086, doi:10.1038/4601085a (2009).
- 99 Rowland, B. D., Bernards, R. & Peeper, D. S. The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. *Nature cell biology* **7**, 1074-1082, doi:10.1038/ncb1314 (2005).
- 100 Tanaka, Y., Era, T., Nishikawa, S. & Kawamata, S. Forced expression of Nanog in hematopoietic stem cells results in a gammadeltaT-cell disorder. *Blood* **110**, 107-115, doi:10.1182/blood-2006-08-039628 (2007).
- 101 Chen, Y. *et al.* The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer. *The Journal of biological chemistry* **283**, 17969-17978, doi:10.1074/jbc.M802917200 (2008).
- 102 Matsuoka, J. *et al.* Role of the stemness factors sox2, oct3/4, and nanog in gastric carcinoma. *The Journal of surgical research* **174**, 130-135, doi:10.1016/j.jss.2010.11.903 (2012).
- 103 Campos-Sanchez, E. & Cobaleda, C. Tumoral reprogramming: Plasticity takes a walk on the wild side. *Biochimica et biophysica acta* **1849**, 436-447, doi:10.1016/j.bbagrm.2014.07.003 (2015).
- 104 Papp, B. & Plath, K. Epigenetics of reprogramming to induced pluripotency. *Cell* **152**, 1324-1343, doi:10.1016/j.cell.2013.02.043 (2013).

- 105 Riggi, N. *et al.* EWS-FLI-1 modulates miRNA145 and SOX2 expression to initiate mesenchymal stem cell reprogramming toward Ewing sarcoma cancer stem cells. *Genes & development* **24**, 916-932, doi:10.1101/gad.1899710 (2010).
- 106 Garcia, C. B. *et al.* Reprogramming of mesenchymal stem cells by the synovial sarcoma-associated oncogene SYT-SSX2. *Oncogene* **31**, 2323-2334, doi:10.1038/onc.2011.418 (2012).
- 107 World Health Organization: Geneva, S. ICD-10: international statistical classification of diseases and related health problems. Tenth Revision. (1992).
- 108 Jaffe ES, H. N., Stein H, Vardiman JW. Pathology and genetics of tumours of haematopoietic and lymphoid tissues. *Lyon: International Agency for Research on Cancer* (2001).
- 109 Muller, A. M., Ihorst, G., Mertelsmann, R. & Engelhardt, M. Epidemiology of non-Hodgkin's lymphoma (NHL): trends, geographic distribution, and etiology. *Annals of hematology* **84**, 1-12, doi:10.1007/s00277-004-0939-7 (2005).
- 110 Anderson, J. R., Armitage, J. O. & Weisenburger, D. D. Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma Classification Project. *Annals of oncology : official journal of the European Society for Medical Oncology* **9**, 717-720 (1998).
- 111 Montoto, S. *et al.* Risk and clinical implications of transformation of follicular lymphoma to diffuse large B-cell lymphoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **25**, 2426-2433, doi:10.1200/JCO.2006.09.3260 (2007).
- 112 Al-Tourah, A. J. *et al.* Population-based analysis of incidence and outcome of transformed non-Hodgkin's lymphoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **26**, 5165-5169, doi:10.1200/JCO.2008.16.0283 (2008).
- 113 Kuppers, R., Klein, U., Hansmann, M. L. & Rajewsky, K. Cellular origin of human B-cell lymphomas. *The New England journal of medicine* **341**, 1520-1529, doi:10.1056/NEJM19991113412007 (1999).
- 114 Tsujimoto, Y., Cossman, J., Jaffe, E. & Croce, C. M. Involvement of the bcl-2 gene in human follicular lymphoma. *Science* **228**, 1440-1443 (1985).
- 115 Yunis, J. J. *et al.* Distinctive chromosomal abnormalities in histologic subtypes of non-Hodgkin's lymphoma. *The New England journal of medicine* **307**, 1231-1236, doi:10.1056/NEJM19821113072002 (1982).
- 116 Weiss, L. M., Warnke, R. A., Sklar, J. & Cleary, M. L. Molecular analysis of the t(14;18) chromosomal translocation in malignant lymphomas. *The New England journal of medicine* **317**, 1185-1189, doi:10.1056/NEJM198711053171904 (1987).
- 117 Horsman, D. E., Gascoyne, R. D., Coupland, R. W., Coldman, A. J. & Adomat, S. A. Comparison of cytogenetic analysis, southern analysis,

- and polymerase chain reaction for the detection of t(14; 18) in follicular lymphoma. *American journal of clinical pathology* **103**, 472-478 (1995).
- 118 Limpens, J. *et al.* Lymphoma-associated translocation t(14;18) in blood B cells of normal individuals. *Blood* **85**, 2528-2536 (1995).
- 119 Roulland, S. *et al.* Long-term clonal persistence and evolution of t(14;18)-bearing B cells in healthy individuals. *Leukemia* **20**, 158-162, doi:10.1038/sj.leu.2404035 (2006).
- 120 Strasser, A., Harris, A. W., Bath, M. L. & Cory, S. Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. *Nature* **348**, 331-333, doi:10.1038/348331a0 (1990).
- 121 McDonnell, T. J. & Korsmeyer, S. J. Progression from lymphoid hyperplasia to high-grade malignant lymphoma in mice transgenic for the t(14; 18). *Nature* **349**, 254-256, doi:10.1038/349254a0 (1991).
- 122 Horsman, D. E. *et al.* Follicular lymphoma lacking the t(14;18)(q32;q21): identification of two disease subtypes. *British journal of haematology* **120**, 424-433 (2003).
- 123 Skibola, C. F. *et al.* Genome-wide association study identifies five susceptibility loci for follicular lymphoma outside the HLA region. *American journal of human genetics* **95**, 462-471, doi:10.1016/j.ajhg.2014.09.004 (2014).
- 124 Horsman, D. E., Connors, J. M., Pantzar, T. & Gascoyne, R. D. Analysis of secondary chromosomal alterations in 165 cases of follicular lymphoma with t(14;18). *Genes, chromosomes & cancer* **30**, 375-382, doi:10.1002/gcc.1103 (2001).
- 125 Johnson, N. A. *et al.* Prognostic significance of secondary cytogenetic alterations in follicular lymphomas. *Genes, chromosomes & cancer* **47**, 1038-1048, doi:10.1002/gcc.20606 (2008).
- 126 Husson, H. *et al.* Gene expression profiling of follicular lymphoma and normal germinal center B cells using cDNA arrays. *Blood* **99**, 282-289 (2002).
- 127 Glas, A. M. *et al.* Gene expression profiling in follicular lymphoma to assess clinical aggressiveness and to guide the choice of treatment. *Blood* **105**, 301-307, doi:10.1182/blood-2004-06-2298 (2005).
- 128 Roehle, A. *et al.* MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas. *British journal of haematology* **142**, 732-744, doi:10.1111/j.1365-2141.2008.07237.x (2008).
- 129 Wang, W. *et al.* MicroRNA profiling of follicular lymphoma identifies microRNAs related to cell proliferation and tumor response. *Haematologica* **97**, 586-594, doi:10.3324/haematol.2011.048132 (2012).
- 130 Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature genetics* **42**, 181-185, doi:10.1038/ng.518 (2010).

- 131 Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303, doi:10.1038/nature10351 (2011).
- 132 Pasqualucci, L. *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* **471**, 189-195, doi:10.1038/nature09730 (2011).
- 133 Green, M. R. *et al.* Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood* **121**, 1604-1611, doi:10.1182/blood-2012-09-457283 (2013).
- 134 Okosun, J. *et al.* Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nature genetics* **46**, 176-181, doi:10.1038/ng.2856 (2014).
- 135 Pasqualucci, L. *et al.* Genetics of follicular lymphoma transformation. *Cell reports* **6**, 130-140, doi:10.1016/j.celrep.2013.12.027 (2014).
- 136 Li, H. *et al.* Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. *Blood* **123**, 1487-1498, doi:10.1182/blood-2013-05-500264 (2014).
- 137 Bouska, A. *et al.* Combined copy number and mutation analysis identifies oncogenic pathways associated with transformation of follicular lymphoma. *Leukemia* **31**, 83-91, doi:10.1038/leu.2016.175 (2017).
- 138 Krysiak, K. *et al.* Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. *Blood* **129**, 473-483, doi:10.1182/blood-2016-07-729954 (2017).
- 139 Timp, W. & Feinberg, A. P. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nature reviews. Cancer* **13**, 497-510, doi:10.1038/nrc3486 (2013).
- 140 Pastore, A. *et al.* Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. *The lancet oncology* **16**, 1111-1122, doi:10.1016/S1470-2045(15)00169-2 (2015).
- 141 Schmidt, J. *et al.* Genome-wide analysis of pediatric-type follicular lymphoma reveals low genetic complexity and recurrent alterations of TNFRSF14 gene. *Blood* **128**, 1101-1111, doi:10.1182/blood-2016-03-703819 (2016).
- 142 Louissaint, A., Jr. *et al.* Pediatric-type nodal follicular lymphoma: a biologically distinct lymphoma with frequent MAPK pathway mutations. *Blood* **128**, 1093-1100, doi:10.1182/blood-2015-12-682591 (2016).
- 143 Green, M. R. *et al.* Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E1116-1125, doi:10.1073/pnas.1501199112 (2015).

- 144 Martelli, M. *et al.* Diffuse large B-cell lymphoma. *Critical reviews in oncology/hematology* **87**, 146-171, doi:10.1016/j.critrevonc.2012.12.009 (2013).
- 145 Fisher, S. G. & Fisher, R. I. The epidemiology of non-Hodgkin's lymphoma. *Oncogene* **23**, 6524-6534, doi:10.1038/sj.onc.1207843 (2004).
- 146 S.H Swerdlow, E. C., N.L. Harris, E.S. Jaffe, S.A. Pileri, H. Stein, J. Thiele, J.W. Vardiman. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. *IARC Press* (2008).
- 147 Meyer, P. N. *et al.* Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **29**, 200-207, doi:10.1200/JCO.2010.30.0368 (2011).
- 148 Choi, W. W. *et al.* A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clinical cancer research : an official journal of the American Association for Cancer Research* **15**, 5494-5502, doi:10.1158/1078-0432.CCR-09-0113 (2009).
- 149 Hans, C. P. *et al.* Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **103**, 275-282, doi:10.1182/blood-2003-05-1545 (2004).
- 150 Alizadeh, A. A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503-511, doi:10.1038/35000501 (2000).
- 151 Offit, K. *et al.* Rearrangement of the bcl-6 gene as a prognostic marker in diffuse large-cell lymphoma. *The New England journal of medicine* **331**, 74-80, doi:10.1056/NEJM199407143310202 (1994).
- 152 Ye, B. H. *et al.* The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nature genetics* **16**, 161-170, doi:10.1038/ng0697-161 (1997).
- 153 Ye, B. H. *et al.* Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma. *Science* **262**, 747-750 (1993).
- 154 Ye, B. H. *et al.* Chromosomal translocations cause deregulated BCL6 expression by promoter substitution in B cell lymphoma. *The EMBO journal* **14**, 6209-6217 (1995).
- 155 Phan, R. T. & Dalla-Favera, R. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature* **432**, 635-639, doi:10.1038/nature03147 (2004).
- 156 Volpe, G. *et al.* Molecular heterogeneity of B-lineage diffuse large cell lymphoma. *Genes, chromosomes & cancer* **16**, 21-30, doi:10.1002/(SICI)1098-2264(199605)16:1<21::AID-GCC3>3.0.CO;2-5 (1996).

- 157 Ueda, C., Nishikori, M., Kitawaki, T., Uchiyama, T. & Ohno, H. Coexistent rearrangements of c-MYC, BCL2, and BCL6 genes in a diffuse large B-cell lymphoma. *International journal of hematology* **79**, 52-54 (2004).
- 158 Thieblemont, C. & Briere, J. MYC, BCL2, BCL6 in DLBCL: impact for clinics in the future? *Blood* **121**, 2165-2166, doi:10.1182/blood-2013-01-480392 (2013).
- 159 Lunning, M. A. & Green, M. R. Mutation of chromatin modifiers; an emerging hallmark of germinal center B-cell lymphomas. *Blood cancer journal* **5**, e361, doi:10.1038/bcj.2015.89 (2015).
- 160 Pasqualucci, L. The genetic basis of diffuse large B-cell lymphoma. *Current opinion in hematology* **20**, 336-344, doi:10.1097/MOH.0b013e3283623d7f (2013).
- 161 Chrivia, J. C. *et al.* Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365**, 855-859, doi:10.1038/365855a0 (1993).
- 162 Wydner, K. L., Bhattacharya, S., Eckner, R., Lawrence, J. B. & Livingston, D. M. Localization of human CREB-binding protein gene (CREBBP) to 16p13.2-p13.3 by fluorescence in situ hybridization. *Genomics* **30**, 395-396 (1995).
- 163 Ait-Si-Ali, S. *et al.* Histone acetyltransferase activity of CBP is controlled by cycle-dependent kinases and oncoprotein E1A. *Nature* **396**, 184-186, doi:10.1038/24190 (1998).
- 164 Eckner, R. *et al.* Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes & development* **8**, 869-884 (1994).
- 165 Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes & development* **14**, 1553-1577 (2000).
- 166 Shi, Y. & Mello, C. A CBP/p300 homolog specifies multiple differentiation pathways in *Caenorhabditis elegans*. *Genes & development* **12**, 943-955 (1998).
- 167 Polesskaya, A. *et al.* CBP/p300 and muscle differentiation: no HAT, no muscle. *The EMBO journal* **20**, 6816-6825, doi:10.1093/emboj/20.23.6816 (2001).
- 168 Grossman, S. R. p300/CBP/p53 interaction and regulation of the p53 response. *European journal of biochemistry* **268**, 2773-2778 (2001).
- 169 Attar, N. & Kurdistani, S. K. Exploitation of EP300 and CREBBP Lysine Acetyltransferases by Cancer. *Cold Spring Harbor perspectives in medicine* **7**, doi:10.1101/cshperspect.a026534 (2017).
- 170 Horwitz, G. A. *et al.* Adenovirus small e1a alters global patterns of histone modification. *Science* **321**, 1084-1085, doi:10.1126/science.1155544 (2008).
- 171 Jin, Q. *et al.* Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *The EMBO journal* **30**, 249-262, doi:10.1038/emboj.2010.318 (2011).

- 172 Turnell, A. S. *et al.* The APC/C and CBP/p300 cooperate to regulate transcription and cell-cycle progression. *Nature* **438**, 690-695, doi:10.1038/nature04151 (2005).
- 173 Tanaka, Y. *et al.* Extensive brain hemorrhage and embryonic lethality in a mouse null mutant of CREB-binding protein. *Mechanisms of development* **95**, 133-145 (2000).
- 174 Petrij, F. *et al.* Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* **376**, 348-351, doi:10.1038/376348a0 (1995).
- 175 Kuhlen, M. & Borkhardt, A. Cancer susceptibility syndromes in children in the area of broad clinical use of massive parallel sequencing. *European journal of pediatrics* **174**, 987-997, doi:10.1007/s00431-015-2565-x (2015).
- 176 Roelfsema, J. H. *et al.* Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *American journal of human genetics* **76**, 572-580, doi:10.1086/429130 (2005).
- 177 Tanaka, Y. *et al.* Abnormal skeletal patterning in embryos lacking a single Cbp allele: a partial similarity with Rubinstein-Taybi syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 10215-10220 (1997).
- 178 Rebel, V. I. *et al.* Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14789-14794, doi:10.1073/pnas.232568499 (2002).
- 179 Alarcon, J. M. *et al.* Chromatin acetylation, memory, and LTP are impaired in CBP<sup>+/-</sup> mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. *Neuron* **42**, 947-959, doi:10.1016/j.neuron.2004.05.021 (2004).
- 180 Kishimoto, M. *et al.* Mutations and deletions of the CBP gene in human lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **11**, 512-519 (2005).
- 181 Mullighan, C. G. *et al.* CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* **471**, 235-239, doi:10.1038/nature09727 (2011).
- 182 Morin, R. D. *et al.* Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* **122**, 1256-1265, doi:10.1182/blood-2013-02-483727 (2013).
- 183 Pasqualucci, L. & Dalla-Favera, R. The genetic landscape of diffuse large B-cell lymphoma. *Seminars in hematology* **52**, 67-76, doi:10.1053/j.seminhematol.2015.01.005 (2015).
- 184 Lohr, J. G. *et al.* Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 3879-3884, doi:10.1073/pnas.1121343109 (2012).

- 185 Kridel, R. *et al.* Histological Transformation and Progression in Follicular Lymphoma: A Clonal Evolution Study. *PLoS medicine* **13**, e1002197, doi:10.1371/journal.pmed.1002197 (2016).
- 186 Hunger, S. P. & Mullighan, C. G. Acute Lymphoblastic Leukemia in Children. *The New England journal of medicine* **373**, 1541-1552, doi:10.1056/NEJMra1400972 (2015).
- 187 Karrman, K. & Johansson, B. Pediatric T-cell acute lymphoblastic leukemia. *Genes, chromosomes & cancer* **56**, 89-116, doi:10.1002/gcc.22416 (2017).
- 188 Litzow, M. R. & Ferrando, A. A. How I treat T-cell acute lymphoblastic leukemia in adults. *Blood* **126**, 833-841, doi:10.1182/blood-2014-10-551895 (2015).
- 189 Marks, D. I. & Rowntree, C. Management of adults with T-cell lymphoblastic leukemia. *Blood* **129**, 1134-1142, doi:10.1182/blood-2016-07-692608 (2017).
- 190 Greaves, M. F., Janossy, G., Peto, J. & Kay, H. Immunologically defined subclasses of acute lymphoblastic leukaemia in children: their relationship to presentation features and prognosis. *British journal of haematology* **48**, 179-197 (1981).
- 191 Pui, C. H., Robison, L. L. & Look, A. T. Acute lymphoblastic leukaemia. *Lancet* **371**, 1030-1043, doi:10.1016/S0140-6736(08)60457-2 (2008).
- 192 Sutton, R. *et al.* Persistent MRD before and after allogeneic BMT predicts relapse in children with acute lymphoblastic leukaemia. *British journal of haematology* **168**, 395-404, doi:10.1111/bjh.13142 (2015).
- 193 Belver, L. & Ferrando, A. The genetics and mechanisms of T cell acute lymphoblastic leukaemia. *Nature reviews. Cancer* **16**, 494-507, doi:10.1038/nrc.2016.63 (2016).
- 194 Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-764, doi:10.1038/nature05690 (2007).
- 195 Coustan-Smith, E. *et al.* Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. *The lancet oncology* **10**, 147-156, doi:10.1016/S1470-2045(08)70314-0 (2009).
- 196 Van Vlierberghe, P. *et al.* ETV6 mutations in early immature human T cell leukemias. *The Journal of experimental medicine* **208**, 2571-2579, doi:10.1084/jem.20112239 (2011).
- 197 Zhang, J. *et al.* The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* **481**, 157-163, doi:10.1038/nature10725 (2012).
- 198 Van Vlierberghe, P. *et al.* Prognostic relevance of integrated genetic profiling in adult T-cell acute lymphoblastic leukemia. *Blood* **122**, 74-82, doi:10.1182/blood-2013-03-491092 (2013).

- 199 Niehues, T. *et al.* A classification based on T cell selection-related phenotypes identifies a subgroup of childhood T-ALL with favorable outcome in the COALL studies. *Leukemia* **13**, 614-617 (1999).
- 200 Ferrando, A. A. *et al.* Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. *Cancer cell* **1**, 75-87 (2002).
- 201 Homminga, I. *et al.* Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. *Cancer cell* **19**, 484-497, doi:10.1016/j.ccr.2011.02.008 (2011).
- 202 Graux, C. *et al.* Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia. *Nature genetics* **36**, 1084-1089, doi:10.1038/ng1425 (2004).
- 203 Van Vlierberghe, P. *et al.* PHF6 mutations in T-cell acute lymphoblastic leukemia. *Nature genetics* **42**, 338-342, doi:10.1038/ng.542 (2010).
- 204 Tosello, V. *et al.* WT1 mutations in T-ALL. *Blood* **114**, 1038-1045, doi:10.1182/blood-2008-12-192039 (2009).
- 205 Kleppe, M. *et al.* Deletion of the protein tyrosine phosphatase gene PTPN2 in T-cell acute lymphoblastic leukemia. *Nature genetics* **42**, 530-535, doi:10.1038/ng.587 (2010).
- 206 Begley, C. G. *et al.* Chromosomal translocation in a human leukemic stem-cell line disrupts the T-cell antigen receptor delta-chain diversity region and results in a previously unreported fusion transcript. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 2031-2035 (1989).
- 207 Valge-Archer, V., Forster, A. & Rabbitts, T. H. The LMO1 and LDB1 proteins interact in human T cell acute leukaemia with the chromosomal translocation t(11;14)(p15;q11). *Oncogene* **17**, 3199-3202, doi:10.1038/sj.onc.1202353 (1998).
- 208 Van Vlierberghe, P. *et al.* The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. *Blood* **108**, 3520-3529, doi:10.1182/blood-2006-04-019927 (2006).
- 209 Weng, A. P. *et al.* Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* **306**, 269-271, doi:10.1126/science.1102160 (2004).
- 210 O'Neil, J. *et al.* FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. *The Journal of experimental medicine* **204**, 1813-1824, doi:10.1084/jem.20070876 (2007).
- 211 Batova, A. *et al.* Frequent and selective methylation of p15 and deletion of both p15 and p16 in T-cell acute lymphoblastic leukemia. *Cancer research* **57**, 832-836 (1997).

- 212 Mendes, R. D. *et al.* PTEN microdeletions in T-cell acute lymphoblastic leukemia are caused by illegitimate RAG-mediated recombination events. *Blood* **124**, 567-578, doi:10.1182/blood-2014-03-562751 (2014).
- 213 Enciso, J., Mendoza, L. & Pelayo, R. Normal vs. Malignant hematopoiesis: the complexity of acute leukemia through systems biology. *Frontiers in genetics* **6**, 290, doi:10.3389/fgene.2015.00290 (2015).
- 214 Kakarala, M. & Wicha, M. S. Implications of the cancer stem-cell hypothesis for breast cancer prevention and therapy. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **26**, 2813-2820, doi:10.1200/JCO.2008.16.3931 (2008).
- 215 R, S., N, F. & M, H. Melanoma subpopulations with cancer stem cell phenotypes. *Jonh Wiley & Sons*, 223-234 (2014).
- 216 Shlush, L. I. *et al.* Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506**, 328-333, doi:10.1038/nature13038 (2014).
- 217 Cox, C. V., Diamanti, P., Evely, R. S., Kearns, P. R. & Blair, A. Expression of CD133 on leukemia-initiating cells in childhood ALL. *Blood* **113**, 3287-3296, doi:10.1182/blood-2008-04-154187 (2009).
- 218 Diamanti, P., Cox, C. V., Moppett, J. P. & Blair, A. Parthenolide eliminates leukemia-initiating cell populations and improves survival in xenografts of childhood acute lymphoblastic leukemia. *Blood* **121**, 1384-1393, doi:10.1182/blood-2012-08-448852 (2013).
- 219 T, R. & M, B. Leukemia-iniciating cells in acute lymphoblastic leukemia. *Jonh Wiley & Sons*, 161-170 (2014).
- 220 Rabbitts, T. H. Translocations, master genes, and differences between the origins of acute and chronic leukemias. *Cell* **67**, 641-644 (1991).
- 221 Rabbitts, T. H. & Boehm, T. Structural and functional chimerism results from chromosomal translocation in lymphoid tumors. *Advances in immunology* **50**, 119-146 (1991).
- 222 Lefranc, M. P. & Rabbitts, T. H. Two tandemly organized human genes encoding the T-cell gamma constant-region sequences show multiple rearrangement in different T-cell types. *Nature* **316**, 464-466 (1985).
- 223 Boehm, T., Foroni, L., Kaneko, Y., Perutz, M. F. & Rabbitts, T. H. The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 4367-4371 (1991).
- 224 Royer-Pokora, B., Loos, U. & Ludwig, W. D. TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukaemia with the t(11;14)(p13;q11). *Oncogene* **6**, 1887-1893 (1991).
- 225 Chambers, J. & Rabbitts, T. H. LMO2 at 25 years: a paradigm of chromosomal translocation proteins. *Open Biol* **5**, 150062, doi:10.1098/rsob.150062 (2015).

- 226 Boehm, T., Buluwela, L., Williams, D., White, L. & Rabbitts, T. H. A cluster of chromosome 11p13 translocations found via distinct D-D and D-D-J rearrangements of the human T cell receptor delta chain gene. *The EMBO journal* **7**, 2011-2017 (1988).
- 227 McGuire, E. A. *et al.* The t(11;14)(p15;q11) in a T-cell acute lymphoblastic leukemia cell line activates multiple transcripts, including Ttg-1, a gene encoding a potential zinc finger protein. *Molecular and cellular biology* **9**, 2124-2132 (1989).
- 228 Boehm, T., Foroni, L., Kennedy, M. & Rabbitts, T. H. The rhombotin gene belongs to a class of transcriptional regulators with a potential novel protein dimerisation motif. *Oncogene* **5**, 1103-1105 (1990).
- 229 Foroni, L. *et al.* The rhombotin gene family encode related LIM-domain proteins whose differing expression suggests multiple roles in mouse development. *Journal of molecular biology* **226**, 747-761 (1992).
- 230 Grutz, G., Forster, A. & Rabbitts, T. H. Identification of the LMO4 gene encoding an interaction partner of the LIM-binding protein LDB1/NLI1: a candidate for displacement by LMO proteins in T cell acute leukaemia. *Oncogene* **17**, 2799-2803, doi:10.1038/sj.onc.1202502 (1998).
- 231 Matthews, J. M., Lester, K., Joseph, S. & Curtis, D. J. LIM-domain-only proteins in cancer. *Nature reviews. Cancer* **13**, 111-122, doi:10.1038/nrc3418 (2013).
- 232 Kenny, D. A., Jurata, L. W., Saga, Y. & Gill, G. N. Identification and characterization of LMO4, an LMO gene with a novel pattern of expression during embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 11257-11262 (1998).
- 233 Sanchez-Garcia, I. & Rabbitts, T. H. The LIM domain: a new structural motif found in zinc-finger-like proteins. *Trends in genetics : TIG* **10**, 315-320 (1994).
- 234 Warren, A. J. *et al.* The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. *Cell* **78**, 45-57 (1994).
- 235 Tse, E. *et al.* Null mutation of the Lmo4 gene or a combined null mutation of the Lmo1/Lmo3 genes causes perinatal lethality, and Lmo4 controls neural tube development in mice. *Molecular and cellular biology* **24**, 2063-2073 (2004).
- 236 Cleary, M. L. Oncogenic conversion of transcription factors by chromosomal translocations. *Cell* **66**, 619-622 (1991).
- 237 Sadler, I., Crawford, A. W., Michelsen, J. W. & Beckerle, M. C. Zyxin and cCRP: two interactive LIM domain proteins associated with the cytoskeleton. *The Journal of cell biology* **119**, 1573-1587 (1992).
- 238 Valge-Archer, V. E. *et al.* The LIM protein RBTN2 and the basic helix-loop-helix protein TAL1 are present in a complex in erythroid cells. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 8617-8621 (1994).
- 239 Schmeichel, K. L. & Beckerle, M. C. The LIM domain is a modular protein-binding interface. *Cell* **79**, 211-219 (1994).

- 240 Wadman, I. A. *et al.* The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *The EMBO journal* **16**, 3145-3157, doi:10.1093/emboj/16.11.3145 (1997).
- 241 Schmeichel, K. L. & Beckerle, M. C. Molecular dissection of a LIM domain. *Mol Biol Cell* **8**, 219-230 (1997).
- 242 Lecuyer, E. *et al.* Protein stability and transcription factor complex assembly determined by the SCL-LMO2 interaction. *The Journal of biological chemistry* **282**, 33649-33658, doi:10.1074/jbc.M703939200 (2007).
- 243 Wadman, I. *et al.* Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. *The EMBO journal* **13**, 4831-4839 (1994).
- 244 Osada, H., Grutz, G., Axelson, H., Forster, A. & Rabbitts, T. H. Association of erythroid transcription factors: complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 9585-9589 (1995).
- 245 Ono, Y., Fukuhara, N. & Yoshie, O. TAL1 and LIM-only proteins synergistically induce retinaldehyde dehydrogenase 2 expression in T-cell acute lymphoblastic leukemia by acting as cofactors for GATA3. *Molecular and cellular biology* **18**, 6939-6950 (1998).
- 246 Agulnick, A. D. *et al.* Interactions of the LIM-domain-binding factor Ldb1 with LIM homeodomain proteins. *Nature* **384**, 270-272, doi:10.1038/384270a0 (1996).
- 247 Yamada, Y., Pannell, R., Forster, A. & Rabbitts, T. H. The oncogenic LIM-only transcription factor Lmo2 regulates angiogenesis but not vasculogenesis in mice. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 320-324 (2000).
- 248 Shivdasani, R. A., Mayer, E. L. & Orkin, S. H. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* **373**, 432-434, doi:10.1038/373432a0 (1995).
- 249 Robb, L. *et al.* Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 7075-7079 (1995).
- 250 Pevny, L. *et al.* Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349**, 257-260, doi:10.1038/349257a0 (1991).
- 251 Simon, M. C. *et al.* Rescue of erythroid development in gene targeted GATA-1- mouse embryonic stem cells. *Nature genetics* **1**, 92-98, doi:10.1038/ng0592-92 (1992).
- 252 Tsai, F. Y. *et al.* An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature* **371**, 221-226, doi:10.1038/371221a0 (1994).

- 253 Durand, C. & Dzierzak, E. Embryonic beginnings of adult hematopoietic stem cells. *Haematologica* **90**, 100-108 (2005).
- 254 Medvinsky, A., Rybtsov, S. & Taoudi, S. Embryonic origin of the adult hematopoietic system: advances and questions. *Development* **138**, 1017-1031, doi:10.1242/dev.040998 (2011).
- 255 Yamada, Y. *et al.* The T cell leukemia LIM protein Lmo2 is necessary for adult mouse hematopoiesis. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 3890-3895 (1998).
- 256 Visvader, J. E., Mao, X., Fujiwara, Y., Hahm, K. & Orkin, S. H. The LIM-domain binding protein Ldb1 and its partner LMO2 act as negative regulators of erythroid differentiation. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 13707-13712 (1997).
- 257 Smith, S. *et al.* LIM domain only-2 (LMO2) induces T-cell leukemia by two distinct pathways. *PloS one* **9**, e85883, doi:10.1371/journal.pone.0085883 (2014).
- 258 Wilson, N. K. *et al.* Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell stem cell* **7**, 532-544, doi:10.1016/j.stem.2010.07.016 (2010).
- 259 Riddell, J. *et al.* Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell* **157**, 549-564, doi:10.1016/j.cell.2014.04.006 (2014).
- 260 Yamada, Y., Pannell, R., Forster, A. & Rabbitts, T. H. The LIM-domain protein Lmo2 is a key regulator of tumour angiogenesis: a new anti-angiogenesis drug target. *Oncogene* **21**, 1309-1315, doi:10.1038/sj.onc.1205285 (2002).
- 261 Deleuze, V. *et al.* TAL-1/SCL and its partners E47 and LMO2 up-regulate VE-cadherin expression in endothelial cells. *Molecular and cellular biology* **27**, 2687-2697, doi:10.1128/MCB.00493-06 (2007).
- 262 Hinks, G. L. *et al.* Expression of LIM protein genes Lmo1, Lmo2, and Lmo3 in adult mouse hippocampus and other forebrain regions: differential regulation by seizure activity. *J Neurosci* **17**, 5549-5559 (1997).
- 263 Herblot, S., Steff, A. M., Hugo, P., Aplan, P. D. & Hoang, T. SCL and LMO1 alter thymocyte differentiation: inhibition of E2A-HEB function and pre-T alpha chain expression. *Nature immunology* **1**, 138-144, doi:10.1038/77819 (2000).
- 264 McCormack, M. P., Forster, A., Drynan, L., Pannell, R. & Rabbitts, T. H. The LMO2 T-cell oncogene is activated via chromosomal translocations or retroviral insertion during gene therapy but has no mandatory role in normal T-cell development. *Molecular and cellular biology* **23**, 9003-9013 (2003).

- 265 Liu, Y. *et al.* The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nature genetics* **49**, 1211-1218, doi:10.1038/ng.3909 (2017).
- 266 Larson, R. C. *et al.* T cell tumours of disparate phenotype in mice transgenic for Rbtn-2. *Oncogene* **9**, 3675-3681 (1994).
- 267 Larson, R. C., Osada, H., Larson, T. A., Lavenir, I. & Rabbitts, T. H. The oncogenic LIM protein Rbtn2 causes thymic developmental aberrations that precede malignancy in transgenic mice. *Oncogene* **11**, 853-862 (1995).
- 268 Cleveland, S. M. *et al.* Lmo2 induces hematopoietic stem cell-like features in T-cell progenitor cells prior to leukemia. *Stem cells* **31**, 882-894, doi:10.1002/stem.1345 (2013).
- 269 Ruggero, K. *et al.* LMO2 and IL2RG synergize in thymocytes to mimic the evolution of SCID-X1 gene therapy-associated T-cell leukaemia. *Leukemia* **30**, 1959-1962, doi:10.1038/leu.2016.116 (2016).
- 270 Gutierrez, A. & Look, A. T. NOTCH and PI3K-AKT pathways intertwined. *Cancer cell* **12**, 411-413, doi:10.1016/j.ccr.2007.10.027 (2007).
- 271 Natkunam, Y. *et al.* The oncoprotein LMO2 is expressed in normal germinal-center B cells and in human B-cell lymphomas. *Blood* **109**, 1636-1642, doi:10.1182/blood-2006-08-039024 (2007).
- 272 Cubedo, E. *et al.* Identification of LMO2 transcriptome and interactome in diffuse large B-cell lymphoma. *Blood* **119**, 5478-5491, doi:10.1182/blood-2012-01-403154 (2012).
- 273 Natkunam, Y. *et al.* LMO2 protein expression predicts survival in patients with diffuse large B-cell lymphoma treated with anthracycline-based chemotherapy with and without rituximab. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **26**, 447-454, doi:10.1200/JCO.2007.13.0690 (2008).
- 274 de Boer, J. *et al.* The E2A-HLF oncogenic fusion protein acts through Lmo2 and Bcl-2 to immortalize hematopoietic progenitors. *Leukemia* **25**, 321-330, doi:10.1038/leu.2010.253 (2011).
- 275 Malumbres, R. *et al.* LMO2 expression reflects the different stages of blast maturation and genetic features in B-cell acute lymphoblastic leukemia and predicts clinical outcome. *Haematologica* **96**, 980-986, doi:10.3324/haematol.2011.040568 (2011).
- 276 Deucher, A. M., Qi, Z., Yu, J., George, T. I. & Etzell, J. E. BCL6 expression correlates with the t(1;19) translocation in B-lymphoblastic leukemia. *American journal of clinical pathology* **143**, 547-557, doi:10.1309/AJCPO4U4VYAAOTEL (2015).
- 277 Nakata, K. *et al.* LMO2 is a novel predictive marker for a better prognosis in pancreatic cancer. *Neoplasia* **11**, 712-719 (2009).
- 278 Cavazzana-Calvo, M. *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**, 669-672 (2000).

- 279 Gaspar, H. B. *et al.* Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* **364**, 2181-2187, doi:10.1016/S0140-6736(04)17590-9 (2004).
- 280 Howe, S. J. *et al.* Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *The Journal of clinical investigation* **118**, 3143-3150, doi:10.1172/JCI35798 (2008).
- 281 Hacein-Bey-Abina, S. *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *The Journal of clinical investigation* **118**, 3132-3142, doi:10.1172/JCI35700 (2008).
- 282 Hacein-Bey-Abina, S. *et al.* LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**, 415-419, doi:10.1126/science.1088547 (2003).
- 283 Yamada, K. *et al.* Identification of a high incidence region for retroviral vector integration near exon 1 of the LMO2 locus. *Retrovirology* **6**, 79, doi:10.1186/1742-4690-6-79 (2009).
- 284 Bester, A. C. *et al.* Fragile sites are preferential targets for integrations of MLV vectors in gene therapy. *Gene Ther* **13**, 1057-1059, doi:10.1038/sj.gt.3302752 (2006).
- 285 Nam, C. H. & Rabbitts, T. H. The role of LMO2 in development and in T cell leukemia after chromosomal translocation or retroviral insertion. *Molecular therapy : the journal of the American Society of Gene Therapy* **13**, 15-25, doi:10.1016/j.ymthe.2005.09.010 (2006).
- 286 Dave, U. P. *et al.* Murine leukemias with retroviral insertions at Lmo2 are predictive of the leukemias induced in SCID-X1 patients following retroviral gene therapy. *PLoS genetics* **5**, e1000491, doi:10.1371/journal.pgen.1000491 (2009).
- 287 Pike-Overzet, K. *et al.* Ectopic retroviral expression of LMO2, but not IL2Rgamma, blocks human T-cell development from CD34+ cells: implications for leukemogenesis in gene therapy. *Leukemia* **21**, 754-763, doi:10.1038/sj.leu.2404563 (2007).
- 288 Xu, W. *et al.* Global transcriptional coactivators CREB-binding protein and p300 are highly essential collectively but not individually in peripheral B cells. *Blood* **107**, 4407-4416, doi:10.1182/blood-2005-08-3263 (2006).
- 289 Hobeika, E. *et al.* Testing gene function early in the B cell lineage in mb1-cre mice. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 13789-13794, doi:10.1073/pnas.0605944103 (2006).
- 290 Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E. & Croce, C. M. The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. *Science* **229**, 1390-1393 (1985).
- 291 Jiang, Y. *et al.* CREBBP Inactivation Promotes the Development of HDAC3-Dependent Lymphomas. *Cancer discovery* **7**, 38-53, doi:10.1158/2159-8290.CD-16-0975 (2017).

- 292 Batta, K., Florkowska, M., Kouskoff, V. & Lacaud, G. Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. *Cell reports* **9**, 1871-1884, doi:10.1016/j.celrep.2014.11.002 (2014).
- 293 Mainardi, S. *et al.* Identification of cancer initiating cells in K-Ras driven lung adenocarcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 255-260, doi:10.1073/pnas.1320383110 (2014).
- 294 Camarero, I. R. Reprogramación tumoral en neoplasias linfoides. *Tesis doctoral* **Universidad de Salamanca**, doi:<https://http://www.educacion.gob.es/teseo/mostrarRef.do?ref=1127883> (2014).
- 295 Maser, R. S. *et al.* Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. *Nature* **447**, 966-971, doi:10.1038/nature05886 (2007).
- 296 De Keersmaecker, K. *et al.* The TLX1 oncogene drives aneuploidy in T cell transformation. *Nature medicine* **16**, 1321-1327, doi:10.1038/nm.2246 (2010).
- 297 Diccianni, M. B. *et al.* Shortened survival after relapse in T-cell acute lymphoblastic leukemia patients with p16/p15 deletions. *Leukemia research* **21**, 549-558 (1997).
- 298 Gutierrez, A. *et al.* The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia. *Blood* **118**, 4169-4173, doi:10.1182/blood-2010-11-318873 (2011).
- 299 Weng, A. P. *et al.* c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. *Genes & development* **20**, 2096-2109, doi:10.1101/gad.1450406 (2006).
- 300 Palomero, T. *et al.* Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. *Nature medicine* **13**, 1203-1210, doi:10.1038/nm1636 (2007).
- 301 Thompson, B. J. *et al.* The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia. *The Journal of experimental medicine* **204**, 1825-1835, doi:10.1084/jem.20070872 (2007).
- 302 Van Vlierberghe, P. & Ferrando, A. The molecular basis of T cell acute lymphoblastic leukemia. *The Journal of clinical investigation* **122**, 3398-3406, doi:10.1172/JCI61269 (2012).
- 303 Espinosa, L. *et al.* The Notch/Hes1 pathway sustains NF-kappaB activation through CYLD repression in T cell leukemia. *Cancer cell* **18**, 268-281, doi:10.1016/j.ccr.2010.08.006 (2010).
- 304 D'Altri, T., Gonzalez, J., Aifantis, I., Espinosa, L. & Bigas, A. Hes1 expression and CYLD repression are essential events downstream of Notch1 in T-cell leukemia. *Cell cycle* **10**, 1031-1036 (2011).
- 305 Komuro, H. *et al.* p27KIP1 deletions in childhood acute lymphoblastic leukemia. *Neoplasia* **1**, 253-261 (1999).

- 306 Dohda, T. *et al.* Notch signaling induces SKP2 expression and promotes reduction of p27Kip1 in T-cell acute lymphoblastic leukemia cell lines. *Experimental cell research* **313**, 3141-3152, doi:10.1016/j.yexcr.2007.04.027 (2007).
- 307 Kasper, L. H. *et al.* CBP/p300 double null cells reveal effect of coactivator level and diversity on CREB transactivation. *The EMBO journal* **29**, 3660-3672, doi:10.1038/emboj.2010.235 (2010).
- 308 Bedford, D. C. & Brindle, P. K. Is histone acetylation the most important physiological function for CBP and p300? *Aging (Albany NY)* **4**, 247-255, doi:10.18632/aging.100453 (2012).
- 309 Kasper, L. H., Qu, C., Obenaus, J. C., McGoldrick, D. J. & Brindle, P. K. Genome-wide and single-cell analyses reveal a context dependent relationship between CBP recruitment and gene expression. *Nucleic acids research* **42**, 11363-11382, doi:10.1093/nar/gku827 (2014).
- 310 Kieffer-Kwon, K. R. *et al.* Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* **155**, 1507-1520, doi:10.1016/j.cell.2013.11.039 (2013).
- 311 Belton, J. M. *et al.* Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**, 268-276, doi:10.1016/j.ymeth.2012.05.001 (2012).
- 312 Chapuy, B. *et al.* Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer cell* **24**, 777-790, doi:10.1016/j.ccr.2013.11.003 (2013).
- 313 Smith, S. M., Anastasi, J., Cohen, K. S. & Godley, L. A. The impact of MYC expression in lymphoma biology: beyond Burkitt lymphoma. *Blood Cells Mol Dis* **45**, 317-323, doi:10.1016/j.bcnd.2010.08.002 (2010).
- 314 Tapia, G. *et al.* Immunohistochemical detection of MYC protein correlates with MYC gene status in aggressive B cell lymphomas. *Histopathology* **59**, 672-678, doi:10.1111/j.1365-2559.2011.03978.x (2011).
- 315 Valera, A. *et al.* MYC protein expression and genetic alterations have prognostic impact in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. *Haematologica* **98**, 1554-1562, doi:10.3324/haematol.2013.086173 (2013).
- 316 Wolf, D., Rodova, M., Miska, E. A., Calvet, J. P. & Kouzarides, T. Acetylation of beta-catenin by CREB-binding protein (CBP). *The Journal of biological chemistry* **277**, 25562-25567, doi:10.1074/jbc.M201196200 (2002).
- 317 Testoni, M. *et al.* Gains of MYC locus and outcome in patients with diffuse large B-cell lymphoma treated with R-CHOP. *British journal of haematology* **155**, 274-277, doi:10.1111/j.1365-2141.2011.08675.x (2011).
- 318 Zhang, X. *et al.* Coordinated silencing of MYC-mediated miR-29 by HDAC3 and EZH2 as a therapeutic target of histone modification in

- aggressive B-Cell lymphomas. *Cancer cell* **22**, 506-523, doi:10.1016/j.ccr.2012.09.003 (2012).
- 319 Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23-28 (1976).
- 320 Sanchez-Garcia, I. How tumour cell identity is established? *Seminars in cancer biology* **32**, 1-2, doi:10.1016/j.semcancer.2015.04.004 (2015).
- 321 Hacein-Bey-Abina, S. *et al.* A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency. *The New England journal of medicine* **371**, 1407-1417, doi:10.1056/NEJMoa1404588 (2014).
- 322 De Ravin, S. S. *et al.* Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. *Science translational medicine* **8**, 335ra357, doi:10.1126/scitranslmed.aad8856 (2016).
- 323 Casola, S. *et al.* Tracking germinal center B cells expressing germ-line immunoglobulin gamma1 transcripts by conditional gene targeting. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 7396-7401, doi:10.1073/pnas.0602353103 (2006).
- 324 Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-193 (2003).
- 325 Irizarry, R. A. *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucleic acids research* **31**, e15 (2003).
- 326 Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264, doi:10.1093/biostatistics/4.2.249 (2003).
- 327 Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 5116-5121, doi:10.1073/pnas.091062498 (2001).
- 328 Benjamini, Y. & Y, H. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B* **57**, 289-300 (1995).
- 329 R Development Core Team (2011) R: A Language and Environment for Statistical Computing. Vienna, Austria : the R Foundation for Statistical Computing. Available online at <http://www.R-project.org/>.
- 330 Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome biology* **5**, R80, doi:10.1186/gb-2004-5-10-r80 (2004).
- 331 Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids research* **30**, 207-210 (2002).
- 332 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

- 333 Muller, F. J. *et al.* Regulatory networks define phenotypic classes of human stem cell lines. *Nature* **455**, 401-405, doi:10.1038/nature07213 (2008).
- 334 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
- 335 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
- 336 Barozzi, I., Termanini, A., Minucci, S. & Natoli, G. Fish the ChIPs: a pipeline for automated genomic annotation of ChIP-Seq data. *Biology direct* **6**, 51, doi:10.1186/1745-6150-6-51 (2011).
- 337 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
- 338 Lerdrup, M., Johansen, J. V., Agrawal-Singh, S. & Hansen, K. An interactive environment for agile analysis and visualization of ChIP-sequencing data. *Nature structural & molecular biology* **23**, 349-357, doi:10.1038/nsmb.3180 (2016).
- 339 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).
- 340 Speir, M. L. *et al.* The UCSC Genome Browser database: 2016 update. *Nucleic acids research* **44**, D717-725, doi:10.1093/nar/gkv1275 (2016).
- 341 Sharov, A. A. & Ko, M. S. Exhaustive search for over-represented DNA sequence motifs with CisFinder. *DNA research : an international journal for rapid publication of reports on genes and genomes* **16**, 261-273, doi:10.1093/dnares/dsp014 (2009).
- 342 Sharov, A. A., DB, D. & MS, K. CisView: a browser and database of cis-regulatory modules predicted in the mouse genome. *DNA Research* **13**, 123-134.
- 343 Konig, M., Reichel, M., Marschalek, R., Haas, O. A. & Strehl, S. A highly specific and sensitive fluorescence in situ hybridization assay for the detection of t(4;11)(q21;q23) and concurrent submicroscopic deletions in acute leukaemias. *British journal of haematology* **116**, 758-764 (2002).
- 344 Davison, J. M., Morgan, T. W., Hsi, B. L., Xiao, S. & Fletcher, J. A. Subtracted, unique-sequence, in situ hybridization: experimental and diagnostic applications. *The American journal of pathology* **153**, 1401-1409, doi:10.1016/S0002-9440(10)65727-6 (1998).
- 345 H, L. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* (2013).
- 346 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297-1303, doi:10.1101/gr.107524.110 (2010).
- 347 Koboldt, D. C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome research* **22**, 568-576, doi:10.1101/gr.129684.111 (2012).

- 348 Ng, S. B. *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272-276, doi:10.1038/nature08250 (2009).
- 349 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research* **38**, e164, doi:10.1093/nar/gkq603 (2010).
- 350 Pasqualucci, L. *et al.* Analysis of the coding genome of diffuse large B-cell lymphoma. *Nature genetics* **43**, 830-837, doi:10.1038/ng.892 (2011).
- 351 Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213-219, doi:10.1038/nbt.2514 (2013).
- 352 Bartenhagen, C. & Dugas, M. Robust and exact structural variation detection with paired-end and soft-clipped alignments: SoftSV compared with eight algorithms. *Brief Bioinform* **17**, 51-62, doi:10.1093/bib/bbv028 (2016).
- 353 Boeva, V. *et al.* Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* **28**, 423-425, doi:10.1093/bioinformatics/btr670 (2012).