# Análisis de los mecanismos moleculares de la predisposición genética a través del estudio de la leucemia infantil

**Marta Isidro Hernández**

**UNIVERSIDAD DE SALAMANCA – CSIC**

Instituto de Biología Molecular y Celular del Cáncer

Salamanca, 2023

**D. Isidro Sánchez García**, investigador científico del Consejo Superior de Investigaciones Científicas (CSIC) en el Instituto de Biología Molecular y Celular del Cáncer (IBMCC) de Salamanca,

Y **Dña. Carolina Vicente Dueñas**, investigadora Miguel Servet del Instituto de Salud Carlos III (ISCIII) (CPII19/00024) perteneciente al grupo de pediatría clínica del Instituto de Investigación Biomédica de Salamanca (IBSAL),

CERTIFICAN:

Que el trabajo titulado "**Análisis de los mecanismos moleculares de la predisposición genética a través del estudio de la leucemia infantil**", que presenta la doctoranda **Dña. Marta Isidro Hernández**, 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 11 de mayo de 2023.

SANCHEZ GARCIA ISIDRO JAVIER - [redacted]

Firmado digitalmente por SANCHEZ GARCIA ISIDRO JAVIER - [redacted]
Nombre de reconocimiento (DN): c=ES, serialNumber=IDCES-[redacted] C, givenName=ISIDRO JAVIER, sn=SANCHEZ GARCIA, cn=SANCHEZ GARCIA ISIDRO JAVIER - [redacted]
Fecha: 2023.05.11 10:28:40 +02'00'

VICENTE DUEÑAS CAROLINA - [redacted]

Firmado digitalmente por VICENTE DUEÑAS CAROLINA - [redacted]
Nombre de reconocimiento (DN): c=ES, serialNumber=IDCES-[redacted] givenName=CAROLINA, sn=VICENTE DUEÑAS, cn=VICENTE DUEÑAS CAROLINA - [redacted]
Fecha: 2023.05.11 10:46:29 +02'00'

Fdo. Dr. Isidro Sánchez García

Fdo. Dra. Carolina Vicente Dueñas

El trabajo recogido en la presente tesis doctoral ha dado lugar a la elaboración de un artículo científico:

**Marta Isidro-Hernández**, Ana Casado-García, Ninad Oak, Silvia Alemán-Arteaga, Belén Ruiz-Corzo, Jorge Martínez-Cano, Andrea Mayado, Elena G. Sánchez, Oscar Blanco, Mª Luisa Gaspar, Alberto Orfao, Diego Alonso-López, Javier De Las Rivas, Susana Riesco, Pablo Prieto-Matos, África González-Murillo, Francisco Javier García Criado, María Begoña García Cenador, Manuel Ramírez-Orellana, Belén de Andrés, Carolina Vicente-Dueñas, César Cobaleda, Kim E Nichols y Isidro Sánchez-García. *Immune stress suppresses innate immune signaling in preleukemic precursor B cells to cause leukemia.* **Nature Communications**. (Versión revisada enviada).

Durante la realización de este trabajo doctoral, Marta Isidro Hernández ha disfrutado de una ayuda para la contratación predoctoral cofinanciada por el Fondo Social Europeo y la Consejería de Educación de la Junta de Castilla y León (CSI021-19).

La investigación en el grupo de Carolina Vicente Dueñas está parcialmente financiada por los Proyectos "PI17/00167" y "PI22/00379" y el contrato “Miguel Servet” (CPII19/00024-AES 2017-2020), financiados por el Instituto de Salud Carlos III (ISCIII) y cofinanciados por la Unión Europea (FEDER/FSE) "Una manera de hacer Europa"/ "El FSE invierte en tu futuro".

La investigación en el grupo de Isidro Sánchez García está parcialmente financiada por FEDER; por SAF2015-64420-R MINECO/FEDER, UE; por RTI2018-093314-B-I00 MCIU/AEI/FEDER, UE; por PID2021-122185OB-I00 MCIU/AEI/FEDER, UE; por la fundación Unoentrecielmil (Proyecto CUNINA) y por la Fundación Científica de la Asociación Española Contra el Cáncer (PRYCO211305-SANC).

Siempre me han repetido la importancia de ser agradecida con aquellos que dedican su tiempo a los demás. Hacer un doctorado y la consiguiente escritura de la tesis no suele ser un proceso sencillo, por eso quisiera expresar mi agradecimiento a todas las personas que, de alguna manera, me han mostrado su ayuda para que este trabajo sea una realidad.

En primer lugar, me gustaría agradecer a mis directores de tesis, el Dr. Isidro Sánchez García y la Dra. Carolina Vicente Dueñas, la oportunidad de formar parte de su laboratorio y mostrar siempre disposición ante cualquier duda o ayuda requerida, así como por ser un ejemplo de trabajo y sacrificio para lograr los objetivos propuestos.

A todos los miembros del laboratorio 13 con los que he coincidido a lo largo de estos años, por todo el tiempo que hemos pasado juntos, tanto dentro como fuera del CIC. A Sara y a Guille, fuisteis los primeros que conocí y que me guiasteis para empezar mi camino en el 13. A Inés, gracias por tu paciencia para enseñarme tanto en tan poco tiempo, me dio mucha pena tu marcha y se te echa de menos. A Javi, gracias porque desde el primer día me recibiste con los brazos abiertos y siempre estuviste dispuesto a ayudarme. A Ana y a Silvia, “tronchi y monchi”, gracias por la infinita ayuda durante estos años, por vuestras risas diarias, pero sobre todo por la compañía en los momentos más difíciles y no tan bonitos. Sin duda, el día a día ha sido mucho más llevadero y tenéis gran parte de culpa de que haya conseguido terminar la tesis. Tras viajes, excursiones, fiestas, una pandemia y muchas horas de “palique” me siento orgullosa de poder consideraros mis amigas. A Belén, gracias por siempre tener una sonrisa, por tus abrazos inesperados, pero siempre necesarios y por convertirte en tan poco tiempo en alguien imprescindible. A Andrea, gracias por tu simpatía y por estar siempre dispuesta a echar una mano. Muchísimas gracias a todos por vuestra ayuda y por las experiencias compartidas.

A la Dra. Kim E. Nichols y al Dr. Ninad Oak, del Departamento de Oncología y al Dr. Jun J. Yang del Departamento de Ciencias Farmacéuticas, todos del St Judes Children’s Research Hospital, por su colaboración totalmente indispensable.

Al Dr. Francisco Javier García Criado y a la Dra. María Begoña García Cenador, del departamento de Cirugía de la Universidad de Salamanca, por facilitarnos el trabajo con los animales y por su buena disposición en todo momento.

Al Dr. Rafael Jiménez, del Departamento de Fisiología y Farmacología de la Universidad de Salamanca, por sus siempre buenas palabras y consejos durante todos estos años.

Al Dr. Alberto Orfao y a Andrea Mayado, del laboratorio 11 del Centro de Investigación del Cáncer, por su colaboración y ayuda con los experimentos de medición de citoquinas en el suero sanguíneo.

Al Servicio de Citometría de la Universidad de Salamanca, en especial a Mª Luz, siempre dispuesta a ayudarnos con los análisis de sorting que con tanta profesionalidad nos han hecho.

Al Servicio de Bioinformática del Centro de Investigación del Cáncer, en especial al Dr. Diego Alonso López y al Dr. Javier De Las Rivas, por toda su ayuda y paciencia con los análisis bioinformáticos dedicados a nuestros estudios.

Al Servicio de Genómica del Centro de Investigación del Cáncer por el procesamiento de las muestras de los *arrays* de expresión.

Al servicio de Anatomía Patológica del Centro de Investigación del Cáncer y al Hospital Universitario de Salamanca, en especial al Dr. Óscar Blanco, por llevar a cabo el procesado y análisis anatomopatológico de las muestras de ratón.

Al servicio de Secuenciación de la Universidad de Salamanca, en especial al Dr. Manuel Sánchez Hernández, por llevar a cabo la secuenciación de las muestras de ratón para la detección de mutaciones.

Al Dr. M Busslinger del *Research Institute of Molecular Pathology* de Viena, por cedernos el modelo animal *Pax5*$^{+/-}$.

A todos los miembros del Servicio de Experimentación Animal de la Universidad de Salamanca, por todo el esfuerzo realizado, incluso en plena pandemia, facilitando así nuestro trabajo y permitiéndonos seguir con nuestros experimentos.

A todo el personal de apoyo y gestión del Instituto de Biología Molecular y Celular del Cáncer de Salamanca, por su amabilidad y disposición para facilitar mi día a día en el centro.

A todas las personas que he conocido en el Centro de Investigación del Cáncer en este tiempo y que, de una manera u otra, han formado parte de esta experiencia, permitiéndome también compartir con ellos momentos fuera del centro.

Me gustaría dar las gracias a mis amigos de toda la vida, por interesaros siempre por mi trabajo y a la vez permitirme desconectar cuando lo he necesitado. Gracias por vuestro apoyo, vuestra confianza en mí y vuestro tiempo dedicado. Aunque cada uno estemos en una punta de España y nos cueste cuadrar horarios, nos bastan cinco minutos recordando anécdotas para comprobar que seguimos como siempre.

Sobre todo, quisiera agradecer a mi familia su apoyo incondicional, por preocuparse por mí y “mis ratoncitos” y por alegrarse de mis logros. En especial, gracias a mis padres por siempre buscar lo mejor para mí, por todos los sacrificios que han hecho para que haya podido llegar hasta aquí y ser la persona que soy hoy en día. Gracias por vuestro apoyo infinito, vuestra confianza en mí y el cariño. Os quiero.

# ÍNDICE

# ÍNDICE

**ABREVIATURAS** ..... **3**

**INTRODUCCIÓN** ..... **9**

**1. La leucemia linfoblástica aguda de células B.** ..... 9

1.1. Definición y características de la leucemia linfoblástica aguda de células B. ..... 9

1.2. Incidencia de la LLA-B. ..... 10

1.3. Tratamiento de la leucemia. ..... 11

**2. Etiología de la leucemia linfoblástica aguda de tipo B.** ..... 13

2.1. El papel de la infección en el desarrollo de la LLA-B. ..... 14

2.2. El papel del microbioma en la LLA-B. ..... 16

2.3. La LLA-B como proceso multietapa. ..... 17

2.4. El papel de la genética en el desarrollo de la LLA-B. ..... 19

**3. Susceptibilidad genética asociada a la LLA-B.** ..... 22

3.1. Papel de *ETV6* en la LLA-B. ..... 22

3.2. Papel de *PAX5* en la LLA-B. ..... 25

3.3. Papel de la inmunidad innata en la LLA-B. ..... 26

**HIPÓTESIS** ..... **33**

**OBJETIVOS** ..... **39**

**RESULTADOS** ..... **45**

**1. El estrés inmunológico no tiene un papel promotor de la leucemia vinculada a la predisposición genética del gen *ETV6*.** ..... 45

1.1. La exposición a infecciones no tiene un efecto directo en la génesis de la LLA asociada al modelo de ratón *Etv6*$^{R355X}$. ..... 46

1.2. La alteración del microbioma intestinal no promueve el desarrollo de la leucemia en ratones *Etv6*$^{R355X}$ genéticamente predispuestos. ..... 52

**2. El estrés inmune suprime la señalización de la inmunidad innata en las células B precursoras preleucémicas para causar leucemia.** ..... 63

2.1. La alteración del microbioma intestinal promueve el desarrollo de la leucemia en ratones *Pax5*$^{+/-}$, pero no en otras cepas genéticamente predispuestas. ..... 65

2.2. El entrenamiento inmunitario mediante la exposición temprana a las infecciones no previene el desarrollo de la leucemia en ratones *Pax5*$^{+/-}$. ..... 69

2.3. El estrés inmunológico disminuye la expresión de *Myd88* en células preleucémicas *Pax5*$^{+/-}$. ..... 74

2.4. La reducción en la expresión de *Myd88* aumenta la incidencia de LLA-B en ratones *Pax5*$^{+/-}$ expuestos a infecciones comunes. ..... 77

2.5. Las células B precursoras de *Pax5*$^{+/-}$;*Myd88*$^{+/-}$ presentan un perfil transcripcional inflamatorio. ..... 82

**DISCUSIÓN ..... 91**

**1. El estrés inmunológico no tiene un papel promotor de la leucemia vinculada a la predisposición genética del gen *ETV6*.** ..... 91

**2. El estrés inmune suprime la señalización de la inmunidad innata en las células B precursoras preleucémicas para causar leucemia.** ..... 94

**CONCLUSIONES ..... 103**

**MATERIAL Y MÉTODOS ..... 109**

**1. Modelos animales utilizados en el estudio.** ..... 109

**2. Genotipado de ratones.** ..... 111

2.1. Genotipado de los ratones *Pax5*$^{+/-}$. ..... 111

2.2. Genotipado de los ratones *Sca1-ETV6-RUNX1.* ..... 112

2.3. Genotipado de los ratones *Sca1-Lmo2.* ..... 112

2.4. Genotipado de los ratones *Sca1-BCR-ABLp190*. ..... 113

2.5. Genotipado de los ratones *Myd88*. ..... 113

2.6. Genotipado de los ratones *ETV6*$^{R355X}$. ..... 114

**3. Tratamiento con antibióticos.** ..... 115

**4. Análisis celular mediante citometría de flujo.** ..... 115

**5. Análisis histológicos.** ..... 117

6. **Análisis de clonalidad de inmunoglobulinas (recombinación VDJ).** ..... 117

7. **Extracción de ADN.** ..... 118

8. **Extracción de ARN.** ..... 119

9. **Extracción del ADN de heces de ratones.** ..... 119

10. **Secuenciación y análisis del ARN ribosómico 16S.** ..... 119

11. **PCR array de la vía de señalización murina de los TLRs.** ..... 121

12. **PCR cuantitativa en tiempo real de *Myd88* (q-PCR).** ..... 121

13. **Cuantificación de los niveles de citoquinas en el suero.** ..... 122

14. **Cultivo de células pro-B prelucémicas.** ..... 123

15. **Análisis de datos mediante *microarrays* de expresión.** ..... 123

16. **Análisis de enriquecimiento (GSEA).** ..... 124

17. **Secuenciación del genoma completo (WGS).** ..... 124

18. **Secuenciación Sanger.** ..... 125

19. **Análisis estadísticos.** ..... 125

**ANEXOS ..... 131**

**BIBLIOGRAFÍA ..... 141**

# ABREVIATURAS

# ABREVIATURAS

**Abx** antibiótico

**ADN** ácido desoxirribonucleico

**APC** *allophycocyanin*, aloficocianina

**ARN** ácido ribonucleico

**ARNr** ácido ribonucleico ribosomal

**ASV** *amplicon sequence variants*, variantes de secuencia de los amplicones

**BSAP** *B-cell specific activation protein,* proteína específica activadora de células B

**CF** *conventional facility*, instalación convencional

**CNV** *copy number variation,* cambios en el número de copias

**dNTPs** deoxirribonucleótidos

**ETS** *erythroblast26 transforming specific,* transformación de eritroblastos específica

**ETV6** ETS variante 6

**FC** *fold change*, cambio de expresión

**FCS** *fetal calf serum*, suero fetal bovino

**FITC** *fluorescein isotiocyanate*, isotiocianato de fluoresceína

**GSEA** *gene set enrichment analysis,* analisis de enriquecimiento de un conjunto de genes

**HSC** *hematopoietic stem cell,* célula madre hematopoyética

**IFN** interferón

**Ig** inmunoglobulina

**IL** interleuquina

**IRF** *interferon regulator factor,* factor regulador del interferón

**LLA** leucemia linfoblástica aguda

**LLA-B** leucemia linfoblástica aguda de células B

**LLA-T** leucemia linfoblástica aguda de células T

**MACS** *magnetic activated cell sorting*, separación celular activada por magnetismo

**MO** médula ósea

**MPW** *milipore water*

**Myd88** molécula adaptadora de diferenciación mieloide 88

**NL** nódulos linfáticos

**pb** pares de bases

**PBS** *phosphate-buffered saline*, tampón fosfato salino

**PCA** *principal component analysis*, análisis de componentes principales

**PCR** *polymerase chain reaction*, reacción en cadena de la polimerasa

**PE** *phycoerythrin*, ficoeritrina

**PI** *propidium iodide,* yoduro de propidio

**PNT** *pointed N-terminal*

**q-PCR** PCR cuantitativa

**RAG** *recombination activation genes,* genes activadores de la recombinación

**RCLB** *red cell lysis buffer,* buffer de lisis de eritrocitos

**SNC** sistema nervioso central

**SNV** *single nucleotide variation,* mutaciones de un nucleótido único

**SP** sangre periférica

**SPF** *specific-pathogen free,* condiciones libres de patógenos específicos

**TGF-β** *transforming growth factor beta*, factor de crecimiento transformante beta

**TLR** *Toll like receptor,* receptor de tipo toll

**TNF** *tumor necrosis factor,* factor de necrosis tumoral

**VAF** *variant allele fraction,* fracción de variantes alélicas

**VDJ** segmentos V (variable), D (diverso) y J (*joining*, unión)

**WES** *whole exome sequencing*, secuenciación del exoma completo

**WGS** *whole genome sequencing*, secuenciación del genoma completo

**WT** wild-type, silvestre

# INTRODUCCIÓN

# INTRODUCCIÓN

## 1. La leucemia linfoblástica aguda de células B.

### 1.1. Definición y características de la leucemia linfoblástica aguda de células B.

Las leucemias agudas son un grupo de neoplasias clonales de células hematopoyéticas inmaduras. En función del origen de la célula implicada que prolifera de manera descontrolada se puede hablar, principalmente, de leucemias linfoides, mieloides o indiferenciadas[1,2].

La leucemia linfoblástica aguda (LLA) es un tipo de leucemia que se caracteriza por la acumulación de células inmaduras de tipo blástico en la médula ósea, impidiendo así una correcta hematopoyesis. En el caso de la leucemia linfoblástica aguda de células B (LLA-B), son células en estadio pro-B y pre-B las que se acumulan en la médula ósea como resultado de un bloqueo en el proceso de diferenciación de las células B[3]. Dichas células son capaces de invadir otros tejidos como bazo, nódulos linfáticos, hígado o sistema nervioso central (SNC) gracias a su entrada en la circulación a través de la sangre periférica[4]. Esto tiene como consecuencia la incorrecta función de estos órganos provocando un amplio espectro de síntomas y signos, como esplenomegalia, hepatomegalia, neutropenia, anemia, sangrados abundantes, dificultad para cicatrizar, inflamación de los nódulos linfáticos, así como palidez, fiebre, fatiga, infecciones frecuentes o erupciones cutáneas[5].

Existen algunos factores de riesgo que aumentan la probabilidad de padecer esta enfermedad, pero que, sin embargo, no son definitorios para contraerla. Algunos de ellos son el sexo, la edad, la raza o tener un hermano gemelo idéntico que tenga LLA. De este modo, ser hombre, de mayor edad, de raza blanca y con un gemelo idéntico que desarrolla LLA en el primer año de vida, tiene un mayor riesgo de padecer esta enfermedad[2].

## 1.2. Incidencia de la LLA-B.

El Instituto Nacional del Cáncer de Estados Unidos ha estimado en el año 2022 el diagnóstico de 60.650 casos de leucemia en este país, lo que supone un 3,16% del total de casos de cáncer. Asociada a esta incidencia, se estiman unas 24.000 muertes por leucemia en este 2022, siendo ligeramente superior en hombres que en mujeres[6].

En concreto, la LLA supone casi un 11% de los casos de leucemia que se diagnostican, contando con 6.660 casos estimados en 2022 en EEUU, de los cuales 3.740 fueron en hombres y 2.920 en mujeres[6]. Se estima también un 23,4% de mortalidad, lo que implica 1.560 muertes por LLA en este año[6,7].

La distribución del tipo de leucemia varía en función de la edad. Mientras que los adolescentes tienden a las leucemias de tipo adulto, con una frecuencia creciente de tipos mieloides y crónicos[8]; en niños, están dominadas por un fenotipo linfocítico pre-B[8], siendo la LLA el cáncer más común[5]. Además, la LLA supone el 18% de todos los cánceres infantiles y presenta un pico de incidencia entre los 2 y los 5 años de edad[9] (**Figura 1**).

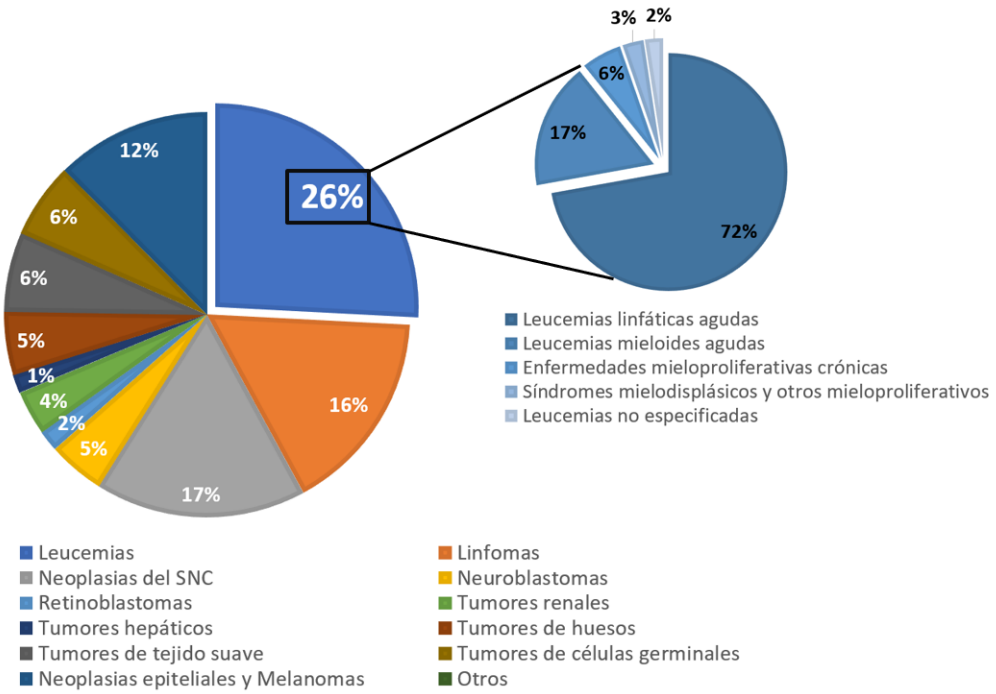



***Figura 1: Incidencia de cánceres pediátricos**. Se muestra el porcentaje de cada tipo de cáncer en la población pediátrica. En el gráfico de la derecha se detalla la distribución de los distintos subtipos de leucemia infantil. Fuente: Surveillance, Epidemiology and End Results Program (SEER) of National Cancer Institute.*

La tasa de incidencia de la leucemia ha aumentado cerca del 1% anual en niños y adolescentes en las dos últimas décadas[7] y se ha mantenido estable en adultos de 20 años o más, aunque las tendencias variaron según el subtipo[7,8]. Este aumento parece estar asociado al estilo de vida moderno, caracterizado por nuevas exposiciones a factores ambientales o circunstancias que afectan principalmente a los niños de los países desarrollados[10]. Además, esto indica que es probable que los factores causales de la enfermedad se hayan vuelto más prevalentes en la población en las últimas décadas[11].

En contraste con la incidencia, la tasa de mortalidad por leucemia ha disminuido en la última década en alrededor del 2% por año en adultos, y en un 3% por año en niños y adolescentes[6,7] (**Figura 2**). Esto se debe en gran parte al avance en el tratamiento de la enfermedad, principalmente a través de la optimización de los agentes quimioterapéuticos establecidos[2] y al desarrollo de fármacos dirigidos[7].

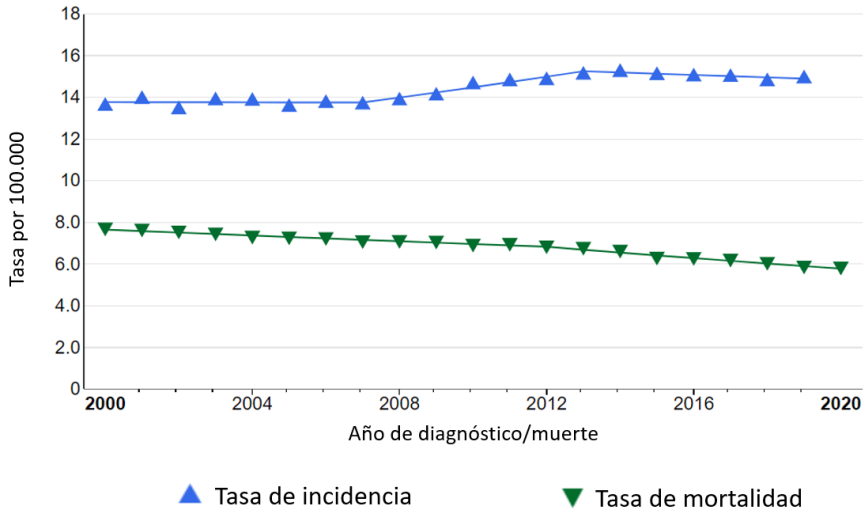



***Figura 2: Incidencia y mortalidad asociadas a la leucemia, ajustadas por años (2000 a 2020).*** *Tasa de incidencia y tasa de mortalidad de la leucemia entre los años 2000 y 2020. Datos teniendo en cuenta ambos sexos, todas las razas y edades. Fuente: Surveillance, Epidemiology and End Results Program (SEER) of National Cancer Institute.*

## 1.3. Tratamiento de la leucemia.

El tratamiento frente a la leucemia suele durar aproximadamente 2 años y comprende tres fases diferenciadas: inducción de la remisión, intensificación (o consolidación) y mantenimiento. Además, las dosis y el calendario de la quimioterapia

combinada se han optimizado en base a las características biológicas de las células leucémicas, la respuesta al tratamiento y los resultados farmacodinámicos y farmacogenómicos de los pacientes, lo que ha dado lugar a una elevada tasa de supervivencia próxima al 90%[12,13,14].

La fase de inducción a la remisión se centra en erradicar la masa de células leucémicas iniciales y restablecer la hematopoyesis normal, generalmente con glucocorticoides, vincristina, daunorrubicina, tratamiento intratecal triple y asparraginasa. Aunque la tasa de remisión tras esta fase es del 96-99% en los niños y del 78-92% en los adultos, es necesario una segunda fase de intensificación en la que se erradican las células leucémicas residuales. Para ello se utilizan varias combinaciones de agentes citotóxicos y la administración de ácido folínico para frenar la toxicidad de la quimioterapia. Finalmente, se entra en una fase de mantenimiento de baja intensidad, que comienza tras recuperar la hematopoyesis y con el paciente en buen estado general, basada en antimetabolitos como mercaptopurina y metotrexato y asparraginasa[12,15].

En las últimas décadas se han logrado tasas de remisión del 90% en la LLA infantil[2], aunque este porcentaje desciende hasta el 40% en el caso de los adultos[7]. Sin embargo, en torno a un 15% de los niños con LLA sufren recaídas, siendo la tasa de curación en estos casos mucho más baja. Hay varios factores que influyen en el pronóstico en el momento de la recaída, como el tiempo transcurrido hasta la recaída (un tiempo más corto se asocia a un peor pronóstico), el inmunofenotipo (el fenotipo de células T se asocia a un peor pronóstico) y el lugar de la recaída (la enfermedad en la médula ósea se asocia a un peor pronóstico que la enfermedad extramedular). Además, las células leucémicas obtenidas de pacientes con recaídas tempranas suelen albergar mutaciones que disminuyen la sensibilidad a los fármacos comunes de la quimioterapia[4]. Para minimizar el riesgo de efectos adversos y recaídas sin reducir la tasa de curación, se requieren terapias más precisas e individualizadas basadas no sólo en la biología de las células leucémicas sino también en las características genómicas del paciente[14,15]. Las inmunoterapias, como las que utilizan anticuerpos biespecíficos o células T modificadas con receptores de antígenos quiméricos, son algunos de los tratamientos prometedores para mejorar el resultado de la LLA resistente[14].

Aun así, las estrategias de prevención son siempre una mejor opción a cualquier tratamiento, por lo que mejoras en el conocimiento de la etiología de la LLA permitirán desarrollar nuevas terapias de prevención que ayuden a la detección precoz e incluso a evitar el desarrollo de la enfermedad.

## 2. Etiología de la leucemia linfoblástica aguda de tipo B.

La leucemia linfoblástica aguda es un claro ejemplo de enfermedad que se origina por la estrecha relación entre la genética y el ambiente. Existen una serie de factores genéticos que confieren una susceptibilidad genética y que en combinación con ciertos factores ambientales van a ser los causantes de la aparición de la LLA.

En cuanto a los factores ambientales, existe una amplia variedad, pero pocos de ellos han podido ser corroborados con datos consistentes o pruebas biológicas. La única exposición causal establecida para la LLA es la radiación ionizante. Esta conclusión se deriva de informes sobre el aumento de la incidencia de la leucemia en los supervivientes de la bomba atómica en Japón[16], en las proximidades de centrales nucleares, así como en niños que han estado expuestos a rayos X durante el embarazo[17]. Estudios epidemiológicos también han sugerido que la exposición a campos electromagnéticos de baja frecuencia por la proximidad a torres de alta tensión o por el cableado doméstico y los aparatos eléctricos son un factor de riesgo de la LLA[16,10,18], pero hasta el momento no hay pruebas biológicas sólidas que lo respalden. Además, los tratamientos de quimioterapia y radioterapia empleados para otros cánceres aumentan el riesgo de sufrir leucemia[19].

Asimismo, hay evidencias que muestran que el alto peso al nacer[20] y el parto mediante cesárea[21] son un determinante del riesgo de la enfermedad. En cualquier caso, hay que tener en cuenta que las asociaciones epidemiológicas no implican una causalidad directa[10]. Por otra parte, la lactancia prolongada durante la infancia ha sido asociada con una leve protección frente al desarrollo de la LLA al proporcionar nutrientes, anticuerpos y moléculas antiinflamatorias. No obstante, no todos los estudios realizados confirman este hecho[22].

Otros posibles agentes ambientales que podrían mediar en el desarrollo de la LLA a pesar de su controversia sobre su participación son: el consumo de alcohol y tabaco en los padres, la dieta de la madre durante el embarazo, la exposición a pesticidas u otros químicos o la contaminación a causa del tráfico[16,23,24].

## 2.1. El papel de la infección en el desarrollo de la LLA-B.

Las infecciones provocadas por cualquier patógeno común son el factor ambiental más estudiado como responsable de la leucemia infantil en pacientes genéticamente predispuestos a esta enfermedad. A pesar de su amplio estudio, ha sido recientemente cuando se ha conseguido establecer una conexión clara entre la exposición a infecciones y el desarrollo de la LLA[10,25,26].

Ya en 1988 fueron propuestas dos hipótesis que postulaban que una respuesta inmunitaria anómala a las infecciones podría ser un importante factor causal del desarrollo de leucemia. La primera hipótesis, "mezcla de poblaciones", fue formulada por Leo Kinlen[27] y defiende el incremento de casos de leucemia infantil en lugares donde ha habido una inusual mezcla de poblaciones, por ejemplo, zonas industriales caracterizadas por una alta migración. Este aumento en la incidencia de la LLA se debe a una respuesta aberrante del sistema inmune a un agente infeccioso de baja patogenicidad[27,28]. La segunda hipótesis, "la exposición tardía", fue propuesta por Mel Greaves, que planteó que una exposición retardada a las infecciones comunes en la infancia genera un sistema inmunitario desregulado, incrementando así el riesgo a desarrollar LLA y manifestando un papel protector de las infecciones tempranas[29,30].

Aunque ambas hipótesis difieren en el supuesto mecanismo que da lugar al desarrollo de la LLA, sí coinciden en que la enfermedad se produce como resultado de una respuesta anormal a las infecciones comunes. De esta manera, la falta de exposición a infecciones en la vida postnatal debido a los "ambientes limpios" de las sociedades desarrolladas podría predisponer al sistema inmune a tener respuestas aberrantes tras la exposición retardada de los niños a infecciones comunes[16].

Varios estudios recientes han aportado pruebas sólidas que confirman la hipótesis de que la exposición a las infecciones puede ser un factor que impulsa la evolución clonal de los clones preleucémicos hacia la leucemia[25,26]. Los datos epidemiológicos apoyan un modelo de "infección tardía" en el que las infecciones comunes promueven los eventos genéticos secundarios posteriores, sólo en el contexto de una exposición infecciosa previa insuficiente en los niños[16]. Este modelo implica que la ausencia de una exposición oportuna a las infecciones en la vida postnatal en los entornos limpios de las sociedades modernas podría impulsar al sistema inmunitario hacia respuestas aberrantes tras una exposición posterior o "retrasada" a patógenos comunes. Sin embargo, también hay estudios que evidencian que las infecciones en el primer año de vida no son protectoras frente a la LLA[31]. En resumen, hay una gran cantidad de datos epidemiológicos que apoyan que el desarrollo de la LLA en niños susceptibles podría ser el resultado de un estrés inmunológico después de la exposición a la infección en la vida postnatal[10].

El uso de modelos animales ha permitido mimetizar un escenario de predisposición a la leucemia como el que se da en humanos, dejando evidencias de que la exposición a la infección representa un entorno oncogénico que promueve el desarrollo de la leucemia. De esta manera, se ha comprobado que modelos murinos heterocigotos para el gen *Pax5* (*Pax5*$^{+/-}$) nunca desarrollan LLA-B en un entorno libre de patógenos. Sin embargo, la exposición de dichos ratones a estímulos infecciosos a la edad de 2 a 5 meses representa el entorno oncogénico necesario para que un 22% acabe desarrollando la enfermedad[25]. También, las LLA-B causadas por la infección que se desarrollan en ratones que expresan la fusión *ETV6-RUNX1* en células madre hematopoyéticas, se asemejan mucho a la enfermedad humana tanto en la baja penetrancia como en la patología y lesiones genómicas. En este modelo, la exposición a la infección también provoca un entorno oncogénico que confiere una presión de selección sobre un compartimento de precursores hematopoyéticos B alterado. Aunque el mecanismo de selección de células tumorales en el modelo *ETV6-RUNX1* es diferente al de los ratones *Pax5*$^{+/-}$, ambos comparten el hecho de que la exposición a las infecciones es un factor desencadenante del desarrollo de una LLA-B completa en ratones genéticamente predispuestos[26].

## 2.2. El papel del microbioma en la LLA-B.

La gran variedad de causas que pueden originar la LLA-B infantil implica que la exposición a la infección no es el único estrés inmunológico capaz de desencadenar la conversión de preleucemia a leucemia[10]. Estudios recientes en ratones *Pax5*$^{+/-}$ y ratones que presentan la fusión *ETV6-RUNX1* muestran que la predisposición genética está sistemáticamente asociada con cambios en el microbioma intestinal, independientemente de que los animales desarrollen o no la LLA-B[32,33]. De esta manera, se ha visto que ratones *Pax5*$^{+/-}$ que han sido expuestos en una fase temprana de su vida a un tratamiento con antibióticos como factor de estrés inmunológico, a pesar de las condiciones de ausencia de patógenos en las que se encontraban, desarrollan LLA-B con una frecuencia aún mayor que los ratones expuestos a infecciones naturales y muestran las mismas alteraciones moleculares y latencia de la enfermedad[32]. Estos resultados tienen importantes implicaciones para el uso de antibióticos en niños, especialmente para los que tienen una predisposición a la LLA-B, ya que estos tratamientos modifican el microbioma intestinal[33].

Esto sugiere que tener predisposición genética a la LLA-B confiere un microbioma intestinal distinto que podría ser utilizado como biomarcador para identificar la susceptibilidad a la LLA-B[32], una posibilidad que se ha propuesto para la LLA-B infantil al igual que para otros trastornos inmunológicos[29].

La microbiota intestinal lleva a cabo varias funciones en nuestro organismo, metabólicas, defensivas o tróficas[34]. En cuanto a su función defensiva, compite contra hongos y bacterias con potencial patogénico por los lugares de fijación y por los nutrientes. Además, inhibe el crecimiento de sus competidores mediante la producción de bacteriocinas[35]. La función trófica de la microbiota intestinal incluye la proliferación y diferenciación de las células epiteliales, la estimulación de la actividad motora intestinal así como la regulación del sistema inmunitario y del SNC[36]. Es por esto que la microbiota intestinal afecta al desarrollo del sistema inmunitario adaptativo al estimular el crecimiento de las estructuras linfoides, la diferenciación de las células T y B y el establecimiento de tolerancia inmunológica[37].

El término disbiosis hace referencia a la perturbación persistente de la microbiota y se ha definido como una alteración tanto de la composición, como de la función y distribución de la microbiota[38]. Este desequilibrio microbiano tiene consecuencias directas en el sistema inmunitario y suele ofrecer una ventaja para la aparición y el crecimiento de microorganismos patógenos, estando este estado del microbioma asociado con la enfermedad[39].

Aunque la mayoría de los tratamientos con antibióticos no tienen efectos adversos, los antibióticos pueden causar cambios significativos en el microbioma que tienen consecuencias para la salud a corto y largo plazo[40]. Hace tiempo que se sospecha una asociación entre la exposición temprana a los antibióticos y la patogénesis de ciertas afecciones cada vez más prevalentes, como las infecciones gastrointestinales, el aumento de peso y la obesidad, la enfermedad inflamatoria intestinal y el cáncer colorrectal[41], así como el asma infantil, las alergias y las enfermedades de las vías respiratorias[42]. El tratamiento con antibióticos también reduce la diversidad general de las especies de la microbiota intestinal provocando cambios metabólicos, aumento de la susceptibilidad del intestino a la colonización y desarrollo de la resistencia bacteriana a los antibióticos, lo que da lugar a infecciones difíciles y costosas de tratar[41].

Por último, la exposición a los antibióticos durante la infancia se ha asociado con un retraso en el desarrollo de la microbiota intestinal. En un estudio en niños de menos de 2 años de edad, el retraso en el desarrollo de la microbiota tras el uso de antibióticos fue especialmente pronunciado entre los 6 y 12 meses de edad[43].

## 2.3. La LLA-B como proceso multietapa.

Se ha establecido que la LLA-B es una enfermedad multietapa que se origina como consecuencia de la cooperación de varias alteraciones genéticas que dan lugar al desarrollo de la enfermedad. En primer lugar, se produce una lesión de susceptibilidad en una célula madre hematopoyética o célula progenitora linfoide que origina un clon preleucémico. Este clon va a ver comprometido su desarrollo normal hacia las células B, pero no lo suficiente como para producir un clon maligno por sí mismo. De este modo,

se va a necesitar que el clon preleucémico adquiera mutaciones secundarias que desencadenen el inicio de la LLA-B[10].

La alteración inicial se puede dar tanto en la línea germinal o en la somática y tiene lugar en el desarrollo prenatal (*in utero*). Se ha propuesto que este primer evento oncogénico surja de manera espontánea, como un error en el desarrollo, pero que, sin embargo, tiene un papel dominante durante el desarrollo de la enfermedad. Como resultado de este evento iniciador se origina un clon preleucémico que puede permanecer silente durante mucho tiempo, incluso toda la vida, sin llegar a producirse la conversión a célula leucémica.

De hecho, una de las características biológicas más relevantes de la LLA-B es la existencia de una fase preleucémica silenciosa latente en la que a pesar de que el primer evento está presente, la leucemia no se desarrolla. Por lo tanto, desde esta perspectiva, la LLA-B infantil es un modelo ideal de estudio para obtener conocimientos sobre cómo prevenir esta transición a una célula maligna completa[10]. Esta condición preleucémica, clínicamente silenciosa, está presente hasta en un 5% de la población infantil sana, pero sólo un mínimo porcentaje, menos del 1%, acaba finalmente desarrollando la leucemia completa[44,45].

Esto indica que el evento inicial genera susceptibilidad al desarrollo de la enfermedad y es condición necesaria, pero no suficiente para que surja la leucemia[46] (con alguna excepción, por ejemplo, las translocaciones del gen *KMT2A* en las LLAs, lesión genética que es suficiente para iniciar el desarrollo de la enfermedad)[47]. De esta manera, el paso crucial en la transición de la preleucemia a la LLA-B es la interacción entre la célula preleucémica y un factor de estrés inmunológico que configura el segundo evento[29]. Este inmuno-estrés está causado por distintos factores ambientales, pero todavía sigue siendo una incógnita por qué en ciertos casos la exposición a un cierto ambiente induce la aparición del segundo evento y en otros no[48] (**Figura 3**).

Muchos estudios se han centrado en entender qué eventos secundarios conducen a la transformación leucémica, y varios de ellos han encontrado respuesta en la posible actividad aberrante de la recombinasa RAG (del inglés, *Recombination Activation Genes*) [49]. Esta enzima está codificada por los genes activadores de la recombinación, conocidos

como RAG1 y RAG2, y lleva a cabo un importante papel en el reordenamiento de los genes de las inmunoglobulinas y el receptor de linfocitos T durante el proceso de recombinación V(D)J. La expresión de esta proteína está restringida a las primeras etapas del desarrollo normal de los linfocitos[50], pero también supone una amenaza, ya que en el momento de máxima actividad y expansión del compartimento de células B (que en los seres humanos coincide con el pico de incidencia de la LLA-B entre los 2 y 5 años), la actividad de la proteína RAG puede causar mutaciones no deseadas siendo responsable de muchos de los segundos eventos que desencadenan la leucemia, como por ejemplo deleciones en *PAX5, IKZF1, ERG, CDKN2A, CDKN2B*, etc.[49,51].

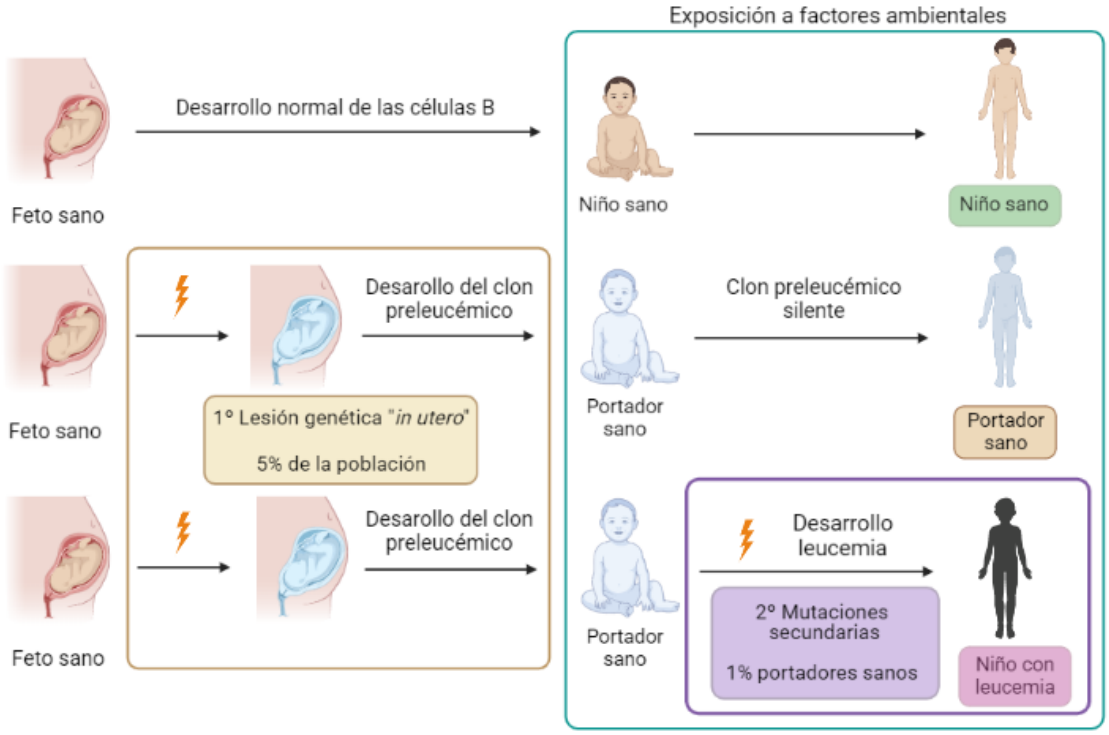



***Figura 3. Desarrollo de la LLA-B como proceso multietapa**. La ausencia de mutaciones durante el desarrollo embrionario da lugar a un niño sano que, aun estando expuesto a factores ambientales, nunca va a desarrollar leucemia. El 5% de la población sufre una primera lesión genética prenatal que genera un clon preleucémico, el cual, en la mayoría de los casos no termina por originar la enfermedad a pesar de la exposición a distintos inmuno-estreses. Un 1% de los portadores sanos del clon preleucémico va a sufrir mutaciones secundarias debido a los factores ambientales, provocando así la aparición de la leucemia.*

## 2.4. El papel de la genética en el desarrollo de la LLA-B.

Gracias a los avances en técnicas de secuenciación y de perfil genómico, el conocimiento de las características moleculares y genómicas de las leucemias ha avanzado mucho en los últimos años[52], permitiendo la clasificación de la mayoría de las

LLAs en distintos subtipos. Estos subtipos de enfermedad están asociados a inicios de edad específicos, a pronósticos variables y a diferentes aberraciones moleculares, de manera que dichas alteraciones iniciadoras de la enfermedad son específicas del inmunofenotipo de la leucemia[53,54].

Las LLAs de fenotipo T se caracterizan por alteraciones en los genes *LMO2*, *TAL1*, *TAL2*, *TLX1*, *TLX2* o *HOXA*; mientras que la hipodiploidía, la hiperdiploidía, la fusión de genes *BCR-ABL1*, *ETV6-RUNX1* y *TCF3-PBX1*, los reordenamientos del gen *KMT2A* o la amplificación intracromosómica del cromosoma 21 son alteraciones propias de la LLA-B[55,56].

Teniendo en cuenta las diferentes alteraciones genéticas iniciadoras, se han podido identificar los siguientes subtipos de LLA-B (**Figura 4A y B**):

- LLAs hipodiploides: se caracterizan por presentar 44 cromosomas o menos. Son más frecuentes en adultos que en niños y en ambos casos se asocia con un mal pronóstico[57].
- LLAs hiperdiploides: se caracterizan por presentar más de 50 cromosomas, tener un buen pronóstico[58] y ser bastante frecuente en niños de entre 2 y 5 años, suponiendo un 21% de los casos de leucemia infantil[56].
- Translocaciones: *ETV6-RUNX1* t(12;21), es la alteración cromosómica más frecuente en la LLA-B infantil y cuenta con un muy buen pronóstico[59]. *BCR-ABL1* t(9;22), es el subtipo más frecuente de LLA-B en adultos. Históricamente asociada a mal pronóstico, el cual mejoró con la aparición de los inhibidores de tirosina quinasas[12]. *TCF3-PBX1* t(1;19), generalmente presenta buen pronóstico, pero está asociada a recaídas que afectan al SNC[12].
- Reordenamientos en el gen *KMT2A*, también conocido como *MLL*: común en leucemias infantiles asociadas a mal pronóstico[60].
- Otros reordenamientos de genes que codifican factores de transcripción como *DUX4*, *MEF2D* y *ZNF384*[15].

Las lesiones iniciadoras más comunes de la LLA-B son las hiperdiploidías, seguidas de los cambios en el número de copias (CNV, del inglés *copy number variations*) como deleciones o translocaciones; y las menos frecuentes son las mutaciones de un

nucleótido (SNV, del inglés *single nucleotide variation*) que se producen en genes clave para el desarrollo hematopoyético y linfoide[29]. Sin embargo, estas alteraciones no son capaces de desencadenar la leucemia por sí solas, sino que necesitan la presencia de mutaciones secundarias que inicien el desarrollo de la enfermedad.

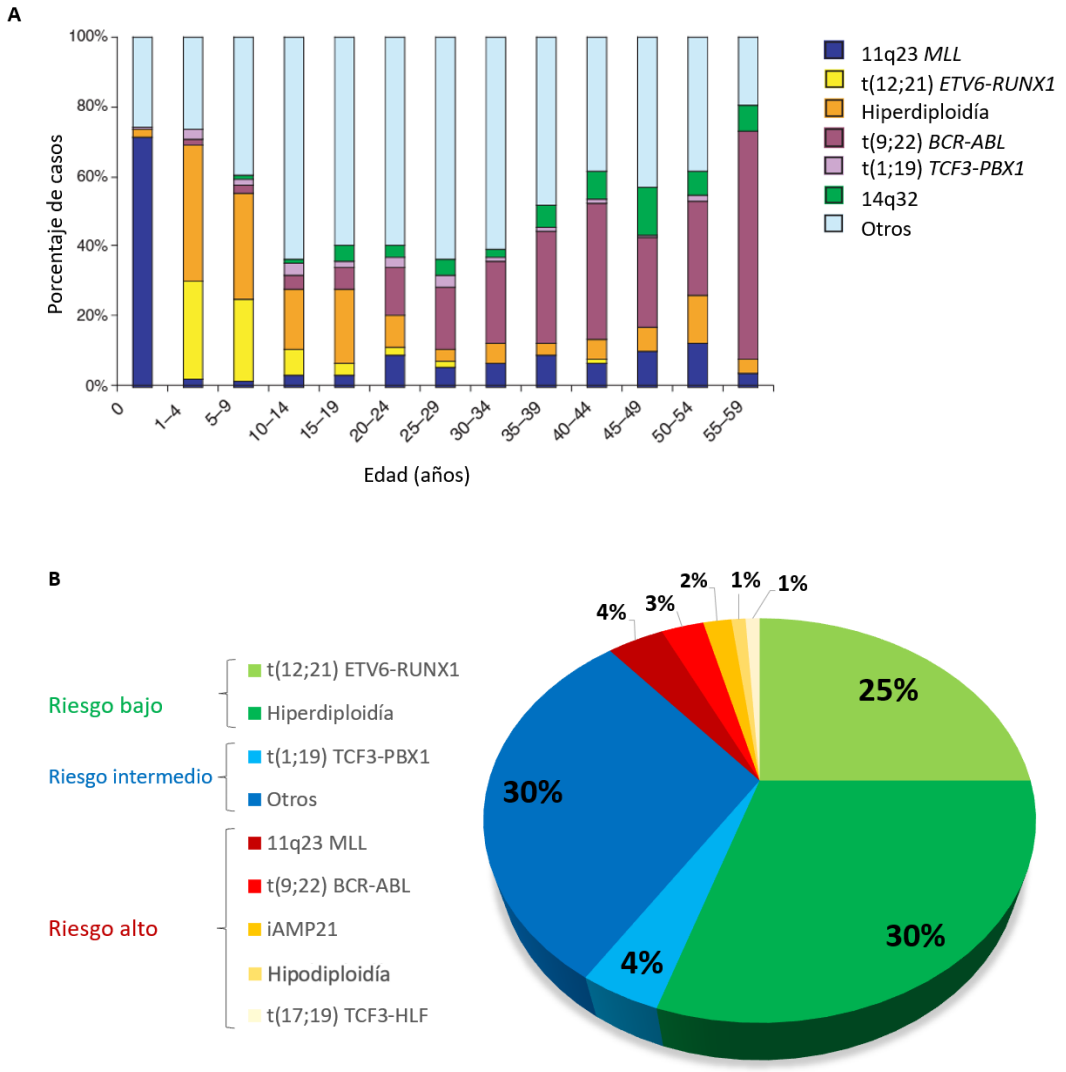



***Figura 4: Distribución de los principales subtipos citogenéticos asociados a la LLA-B. A)*** *Distribución de los subtipos de LLA-B desglosados según la edad. Se observan cambios en la distribución de las distintas alteraciones genéticas según el rango de edad de diagnóstico del paciente con leucemia. Adaptado de Harrison, C.J. British Journal of Haematology 2009.* ***B)*** *Frecuencia de los principales subgrupos citogenéticos de LLA-B y el riesgo asociado: los grupos de riesgo bajo se muestran en verde; los grupos de riesgo intermedio se representan en azul y el rojo indica riesgo alto[61]. Adaptado de Schwab, C. & Harrison, C.J. Hemasphere 2018.*

Entre los genes que aparecen mutados de forma recurrente, se encuentran genes involucrados en importantes funciones biológicas, como: factores de transcripción (*IKZF1*, *PAX5*, *EBF1*, *ETV6*); supresores tumorales y reguladores del ciclo celular (*TP53*, *PTEN*, *RB1*, *CDKN2A*/*CDKN2B*); señalización de receptores de citoquinas (*CRLF2*, *IL7R*,

*EPOR*), quinasas (*ABL1*, *ABL2*, *JAK2*) y Ras (*KRAS*, *NRAS*, *PTPN11*); y modificadores epigenéticos (*EZH2*, *CREBBP*, *SETD2*)[62]. Estudios de secuenciación genómica han determinado que las alteraciones secundarias más comunes ocurren en genes que codifican para factores de transcripción involucrados en la diferenciación de linfocitos B (por ejemplo: *IKZF1*, *PAX5* o *EBF1*), apareciendo en aproximadamente un 40% de las LLA-B[63]. Aunque también se han identificado mutaciones germinales en *PAX5*, *ETV6* e *IKZF1*[15] que predisponen a desarrollar LLA-B y en este caso sí que se tratarían de un evento iniciador. En términos generales, las fusiones de genes tienden a actuar como eventos iniciadores durante la transformación leucémica temprana, mientras que las pequeñas aberraciones genómicas ocurren más a menudo como eventos secundarios tardíos[10].

# 3. Susceptibilidad genética asociada a la LLA-B.

## 3.1. Papel de *ETV6* en la LLA-B.

*ETV6* (ETS *Variant 6*) es un miembro de la familia de factores de transcripción ETS (*Erythroblast26 Transforming Specific*) que están implicados en variedad de procesos de desarrollo y diferenciación celular[64]. El gen humano *ETV6* abarca una región de más de 240 kb en el brazo corto (p) del cromosoma 12 en la banda 13.1 y consta de 8 exones[65].

Las proteínas ETS tienen tres dominios: el dominio ETS, ampliamente conservado y que media la unión de esta proteína al ADN[66]. El dominio *Pointed* (PNT, del inglés *Pointed N-Terminal*), también conservado evolutivamente y responsable de la homo y heterodimerización, mediando así las interacciones proteína-proteína[67]. Y el dominio central, menos conservado y que contribuye a la actividad de represión de *ETV6*[68].

*ETV6* desempeña un papel importante en la hematopoyesis embrionaria, participando en la regulación de las poblaciones eritroides, mieloides y linfoides[69]. La ausencia completa del gen *ETV6* en ratones provoca la letalidad embrionaria debido a la apoptosis de las células mesenquimales y neurales, acompañada de una angiogénesis defectuosa[70]. En adultos, *ETV6* es necesario para la supervivencia de las células madre

hematopoyéticas (HSCs, del inglés *hematopoietic stem cell*) y el establecimiento de la hematopoyesis en la médula ósea[71].

El gen *ETV6* destaca por su frecuente implicación en translocaciones cromosómicas asociadas a neoplasias hematológicas, habiéndose descrito aproximadamente 50 translocaciones que implican al gen *ETV6* junto a otros 30 genes asociados[72,73]. Además de participar en translocaciones que dan lugar a la formación de genes de fusión o a la activación de protooncogenes, *ETV6* funciona como gen supresor de tumores[74].

Distintas alteraciones en *ETV6* están asociadas a diferentes enfermedades, pero en general, las mutaciones somáticas de *ETV6* se asocian a síndromes mielodisplásicos y leucemias de células T. Sin embargo, mutaciones en la línea germinal están relacionadas con la trombocitopenia hereditaria y la susceptibilidad a la LLA-B[75].

Estudios recientes han descrito mutaciones en *ETV6* en la línea germinal en familias que presentan trombocitopenia autosómica dominante y predisposición a la LLA. Estas alteraciones están ausentes en la población general y se asocian a la trombocitopenia en todos los casos. Además, estas variantes disminuyen la actividad represora transcripcional de *ETV6* y la mayoría deterioran su unión al ADN y alteran su localización subcelular[76,77].

Otros estudios han identificado recientemente casos de leucemia familiar con alteraciones estructurales heredadas en el gen *ETV6*, lo que pone de manifiesto el papel de este gen en la predisposición a la LLA[78,79].

El dominio ETS media la unión de *ETV6* al ADN, por lo que su papel es fundamental en la actividad represora transcripcional de dicha proteína. De hecho, mutaciones en el dominio ETS inducen una pérdida casi completa de la actividad represora. Un ejemplo que ilustra este hecho se vio al examinar células de la médula ósea de pacientes portadores de la mutación sin sentido p.R359X y no detectar la proteína truncada, sugiriendo que esta variante da lugar a la pérdida de función de *ETV6*[80].

Por otro lado, el dominio PNT facilita la homodimerización y el dominio central está implicado en la formación de complejos correpresores[67,81]. Por lo tanto, las variantes de *ETV6* en estos dominios o cerca de ellos podrían interrumpir las interacciones proteína-proteína y así alterar la unión de los complejos correpresores a dianas específicas[80].

Dado que las mutaciones en la línea germinal de *ETV6* no siempre dan lugar a malignidad, ya que sólo entre el 25% y el 30% de los portadores desarrollan leucemia[82], es necesario la adquisición de alteraciones genómicas adicionales de forma somática para inducir la leucemia. De hecho, en las LLAs asociadas a *ETV6*, el programa transcripcional está fuertemente influenciado por las anomalías somáticas secundarias, con una separación especialmente llamativa de los casos de LLA hiperdiploide frente a los diploides. En contraste con los casos hiperdiploides caracterizados por mutaciones en la vía de *RAS*, los casos de LLA diploide con variantes de *ETV6* en la línea germinal, muestran una deleción altamente recurrente de *ETV6* y *PAX5* (y en menor grado de *IKZF1*). Estas deleciones de *PAX5* y *ETV6* están ausentes en gran medida en las LLAs hiperdiploides, pero se observan con frecuencia en la LLA que alberga la fusión somática *ETV6-RUNX1* (**Figura 5**). Esto sugiere que, incluso con el mismo factor de predisposición genética subyacente, las células preleucémicas pueden activar unas pocas vías restringidas pero distintivas como eventos principales de cooperación durante la formación de la leucemia[80].

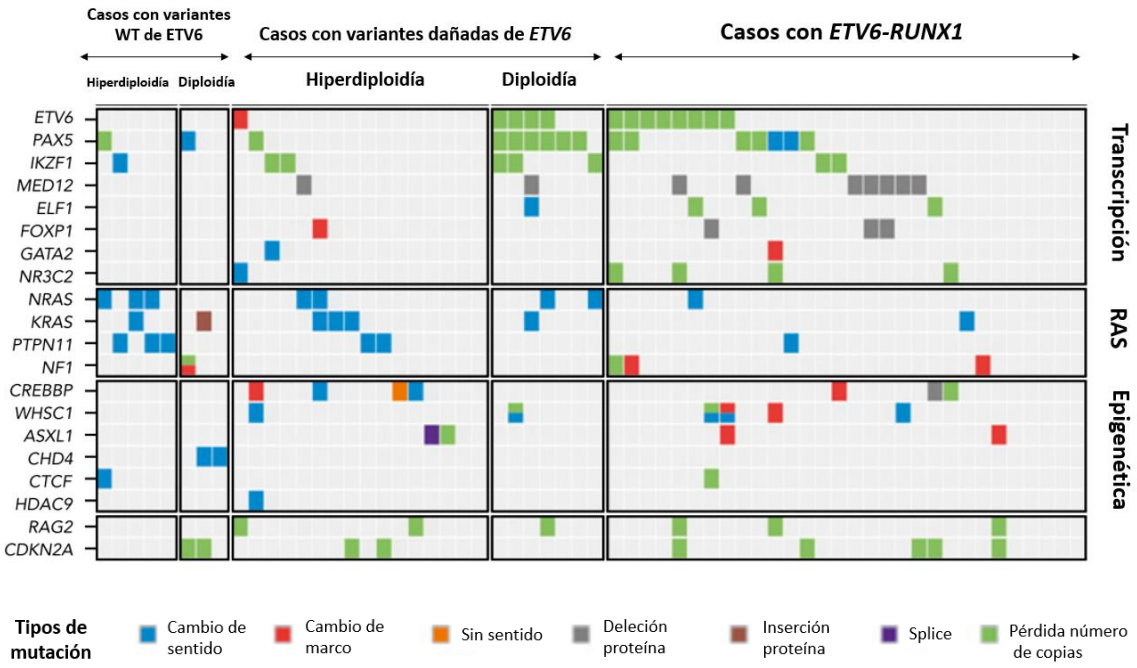



***Figura 5. Perfil genómico completo de la LLA con variantes de ETV6 en la línea germinal****. Heatmap que muestra las principales mutaciones codificantes somáticas identificadas en casos con una variante ETV6 de línea germinal o fusión ETV6-RUNX1. Adaptado de Jun J. Yang et al. Blood, 2021*[80].

De modo que, dado que la hiperdiploidía puede surgir en el útero, esta anomalía genómica temprana podría promover la aparición de la leucemia junto con los defectos de *ETV6* en la línea germinal, con un pequeño número de eventos somáticos adicionales (por ejemplo, mutaciones en *RAS*). Por el contrario, la LLA diploide *ETV6* puede tomar un camino más estrecho hacia la leucemia, requiriendo la pérdida de varios factores de transcripción hematopoyéticos (por ejemplo, *ETV6*, *PAX5* o *IKZF1*) que darían lugar a la interrupción de la diferenciación normal de las células B[80].

## 3.2. Papel de *PAX5* en la LLA-B.

El gen *PAX5* se expresa en células hematopoyéticas, estando su expresión restringida a las células B. *PAX5* codifica la proteína específica activadora de células B (BSAP, del inglés *B-cell Specific Activation Protein*), cuya expresión se extiende desde el estadio pro-B hasta células B maduras, apagándose entonces para permitir la diferenciación de célula B madura a célula plasmática.[83] De esta manera, *PAX5* controla la identidad y el desarrollo de las células B mediante la represión de genes inapropiados para el linaje y la activación de genes específicos de células B[84]. Además, este factor de transcripción ha sido definido como el guardián de la identidad B debido a que, en su ausencia, el desarrollo de células B no es posible y se puede inducir la transdiferenciación a células T o células mieloides aplicando los factores de crecimiento adecuados[83,85].

Trabajos previos avalan el uso del modelo de ratón *Pax5*$^{+/-}$ para el estudio del desarrollo de la LLA-B[25,86], habiéndose demostrado que dichos ratones *Pax5*$^{+/-}$ presentan una acumulación de células proB y preB en la médula ósea y por tanto el desarrollo B está alterado debido a la inactivación de uno de los alelos del gen[25]. De esta manera, los ratones *Pax5*$^{-/-}$ carecen de células B maduras y la expresión de *Cd19* está apagada en los mismos, ya que *Cd19* es una diana directa de *Pax5* que sirve como marcador de actividad del gen.

Como se ha comentado anteriormente, algunas de las alteraciones secundarias más comunes en la LLA-B tienen lugar en factores de transcripción implicados en el desarrollo de células B, siendo *PAX5* uno de los genes que más frecuentemente aparece mutado o

delecionado. De hecho, *PAX5* tiene un rol de supresor tumoral en el desarrollo de la LLA-B, apareciendo alterado en el 30% de las LLA-B[63]. Sin embargo, *PAX5* también ha sido identificado como gen que confiere susceptibilidad al desarrollo de la LLA-B en individuos que presentan mutaciones en la línea germinal[87].

Los modelos murinos han permitido establecer la infección como detonante de la leucemia en ratones *Pax5*$^{+/-}$ que presentan una susceptibilidad a la LLA-B[25]. De igual manera, hay hallazgos epidemiológicos que demuestran que la infección puede actuar como desencadenante del desarrollo de la LLA-B en humanos[29]. Sin embargo, todavía no se sabe cómo este estrés inmunológico desata el potencial de transformación de las células preleucémicas, que de otro modo estaría restringido.

## 3.3. Papel de la inmunidad innata en la LLA-B.

La inmunidad innata desempeña un papel fundamental en la evolución de aquellos cánceres en los que existe una asociación causal con la infección. Se ha descrito que, en los cánceres de causa infecciosa conocida como el cáncer de estómago, de páncreas, de hígado, o de cuello de útero, los receptores de tipo *toll* (TLR, del inglés *Toll-Like Receptor*) desempeñan un papel central en la modulación de las respuestas celulares inhibidoras y promotoras de tumores, tanto en las células inmunes del microambiente tumoral como directamente en las células cancerosas[88]. Por este motivo, podría ser interesante centrarse en la vía de la inmunidad innata mediada por la señalización TLR como método para restaurar la respuesta inmunitaria normal en ratones *Pax5*$^{+/-}$.

Los TLRs están presentes en casi todos los tipos de células y actúan como iniciadores de la inflamación y la inmunidad a través de la inducción de citoquinas proinflamatorias, interferones (IFN) y quimiocinas mediadas por NF-κB y factores reguladores de interferones (IRF, de las siglas en inglés *interferon regulator factor*)[89]. Los TLRs también desempeñan un papel importante en controlar el equilibrio del hospedador con la microbiota intestinal y son capaces de detectar alteraciones en la homeostasis de los tejidos y de activar mecanismos de curación de heridas y de reparación de tejidos[90].

Asimismo, están implicados en el desarrollo del linaje B, el cambio de clase de anticuerpos y la activación y mantenimiento de las células B[91,92].

Los TLRs reconocen ligandos microbianos o endógenos y presentan una señalización intracelular a través de la molécula adaptadora de diferenciación mieloide 88 (*Myd88*), con la excepción de los TLR3 que lo hacen mediante la vía TRIF[93].

*Myd88* actúa en la interfaz de la homeostasis tisular y el cáncer. Por un lado, se ha visto que *Myd88* contribuye en la carcinogénesis en modelos de cáncer de piel, hígado, páncreas y colon, influyendo en la activación de la inflamación mediada por las células hematopoyéticas infiltradas[93]. Sin embargo, en otros modelos, *Myd88* tiene un efecto protector, ya que la ausencia de señalización de *Myd88* en células del microambiente tumoral aumenta el ataque inmunológico a estas células[94]. De igual modo, ratones *Myd88*$^{-/-}$ muestran alteraciones o retrasos en la cicatrización de heridas en la piel[95], en la reparación del tejido del colon[96] y en la regeneración del hígado[97].

*Myd88*, además de inducir la respuesta innata, tiene un papel importante en la maduración de las células B, especialmente al inicio del desarrollo B y en la formación de los centros germinales[98]. Existen antecedentes de una colaboración entre *PAX5* y *Myd88* en enfermedades hematológicas malignas como el linfoma difuso de células B grandes, el linfoma primario del sistema nervioso central o la macroglobulinemia de Waldenstrom; y aunque *Myd88* no induce directamente la expresión de *PAX5* sí que interviene en la regulación de su activación[99].

# HIPÓTESIS

# HIPÓTESIS

La LLA-B es el cáncer infantil más frecuente en el mundo y, a pesar de que un 90% de los casos se curan, sigue siendo una de las principales causas de muerte infantil por enfermedad. Esto se debe a la falta de tratamientos óptimos para las recaídas y a las graves consecuencias a largo plazo que conlleva la toxicidad del tratamiento. Se conoce que al menos el 5% de los niños que nacen son portadores de alteraciones genéticas que participan en el desarrollo de la LLA-B, sin embargo, sólo un pequeño porcentaje termina por desarrollar la enfermedad. Esto conduce a la idea de que el desarrollo de la leucemia sigue un modelo de dos eventos: un primer evento oncogénico durante la etapa prenatal confiere susceptibilidad genética al clon preleucémico que se forma, el cual adquiere mutaciones secundarias por la acción de factores ambientales que conducen a la aparición de la leucemia[10].

Las alteraciones genéticas que conducen a la aparición de un clon preleucémico son múltiples y el segundo evento que finalmente transforma a dicho clon parece dependiente del evento genético inicial[10]. Estudios recientes han descrito mutaciones germinales en el gen *ETV6* en familias que presentan predisposición a la LLA[76,77]. No todos los portadores de la mutación en dicho gen desarrollarán la leucemia, lo que pone de manifiesto, una vez más, que es necesario un evento secundario que dé lugar a la enfermedad, pero dicho evento se desconoce hoy en día. De modo que es posible que el desarrollo de la LLA-B vinculada a la predisposición genética asociada al gen *ETV6* esté mediado por un estrés inmunológico como ocurre con otros genes de susceptibilidad (*PAX5* o *ETV6-RUNX1*). Con el fin de dilucidar esta cuestión, utilizaremos un modelo murino portador de una de las mutaciones encontradas en humanos para explorar el papel del estrés inmunológico en dicho contexto.

Estudios recientes han demostrado que la exposición ambiental a infecciones es el principal causante del estrés inmunitario necesario para proporcionar la ventaja de selección que permite a las células preleucémicas dar lugar a una leucemia completa[25,86]. Sin embargo, todavía no se sabe muy bien cómo este estrés inmunitario desata el potencial de transformación de las células preleucémicas, que de otro modo

estaría restringido. Es por ello por lo que se plantea la idea de que la inmunidad innata a través de la vía de señalización de los TLRs participe en los mecanismos de evolución de las células B preleucémicas en las LLA-B asociadas a la infección.

Para tratar de resolver esta cuestión, se han utilizado los modelos murinos *Pax5*$^{+/-}$ y *Sca1-ETV6-RUNX1,* predispuestos a desarrollar leucemia cuando se someten a un estrés inmunológico como es la exposición a infecciones[25,26], para dilucidar cómo las células preleucémicas cambian a un estado leucémico durante el estrés inmunológico y cuál es el papel de la inmunidad innata en este proceso. El uso de modelos murinos ha facilitado esta labor de investigación debido a que generalmente los primeros pasos en el desarrollo de la LLA-B pasan desapercibidos en los niños, siendo casi imposible estudiar en humanos las etapas iniciales de la transformación de la preleucemia.

# OBJETIVOS

# OBJETIVOS

1. Estudiar el papel del estrés inmunológico en la génesis de la LLA vinculada a la predisposición genética asociada al gen *ETV6*.

2. Estudiar los mecanismos que median el papel del estrés inmunológico en la génesis de la LLA-B.

# RESULTADOS

# RESULTADOS

## 1. El estrés inmunológico no tiene un papel promotor de la leucemia vinculada a la predisposición genética del gen *ETV6*.

Como se ha explicado anteriormente, la LLA-B es una enfermedad multietapa que se origina como consecuencia de la cooperación de varias alteraciones genéticas que dan lugar al desarrollo de la enfermedad. De este modo, un clon preleucémico adquiere una segunda mutación que le permite transformarse en un clon leucémico maligno.

Un ejemplo del modelo de dos eventos tiene lugar con los ratones *Pax5*$^{+/-}$ y *Sca1-ETV6-RUNX1*, los cuales presentan una predisposición genética a desarrollar la enfermedad, pero sin la presencia de una segunda mutación no terminarán desarrollándola. Se ha estudiado que la exposición a infecciones actúa como desencadenante de ese segundo evento, por lo que el traslado de estos modelos de ratones a un ambiente con presencia de patógenos comunes va a provocar que un porcentaje significativo sucumba a la enfermedad debido al estrés inmunológico provocado por la exposición a dichos patógenos[25,26].

Además, recientemente se ha demostrado que la predisposición genética de *Pax5*$^{+/-}$ y *ETV6-RUNX1* conforma un microbioma intestinal único y claramente distinto al de ratones genéticamente no predispuestos a desarrollar LLA-B[32]. Curiosamente, este microbioma intestinal especifico de los ratones *Pax5*$^{+/-}$ protege a estos ratones predispuestos contra el desarrollo de la leucemia, y su alteración mediante el tratamiento con antibióticos desencadena un estrés inmunológico que es suficiente para inducir la LLA-B, incluso en ausencia de estímulos infecciosos[8]. Actualmente, hay muchos trabajos que se están centrando en estudiar la composición del microbioma intestinal y en mostrar así su relevancia en la modulación de ciertas enfermedades como la leucemia.

Estudios recientes han descrito mutaciones germinales en el gen *ETV6* en familias que presentan predisposición a la LLA. Dado que las mutaciones en la línea germinal de *ETV6* no siempre dan lugar a malignidad, ya que sólo un pequeño porcentaje de los portadores desarrollan leucemia[82], es necesario la adquisición de alteraciones genómicas adicionales de forma somática para inducir la leucemia. Esto lleva a pensar que sea posible que el desarrollo de la LLA vinculada a la predisposición genética asociada al gen *ETV6* esté mediado por un estrés inmunológico como ocurre con otros genes de susceptibilidad. Con el fin de dilucidar esta cuestión, en este trabajo se planteó el estudio de un modelo murino portador de una de las mutaciones encontradas en humanos (p.R359X; p.R355X en ratones) para explorar el papel del estrés inmunológico en dicho contexto. Por ello, se quiso someter a ratones *Etv6*$^{R355X}$ predispuestos a desarrollar LLA-B a diferentes estreses inmunológicos para ver si terminaban desarrollando la enfermedad.

## 1.1. La exposición a infecciones no tiene un efecto directo en la génesis de la LLA asociada al modelo de ratón *Etv6*$^{R355X}$.

Para estudiar si la exposición a patógenos comunes induce el desarrollo de la leucemia en los ratones *Etv6*$^{R355X}$ predispuestos, se estableció una cohorte de 34 ratones *Etv6*$^{R355X}$ y 23 ratones de tipo silvestre (WT, del inglés *wild type*). Todos los animales nacieron y se mantuvieron en un animalario con condiciones libres de patógenos específicos (SPF, de sus siglas en inglés *specific-pathogen free*) hasta las 6 semanas de edad, momento en el que fueron trasladados a uno convencional (CF, del inglés *conventional facility*), donde estaban expuestos a infecciones (**Figura 6**). Estos ratones fueron vigilados diariamente para comprobar su estado y se monitorizaron con sangrados rutinarios bimensuales. Además, fueron mantenidos hasta el punto final de experimento (22 meses) o hasta que mostraron signos de enfermedad, momento en el que se sacrificaron para extraer y analizar los distintos órganos (**Figura 6**).

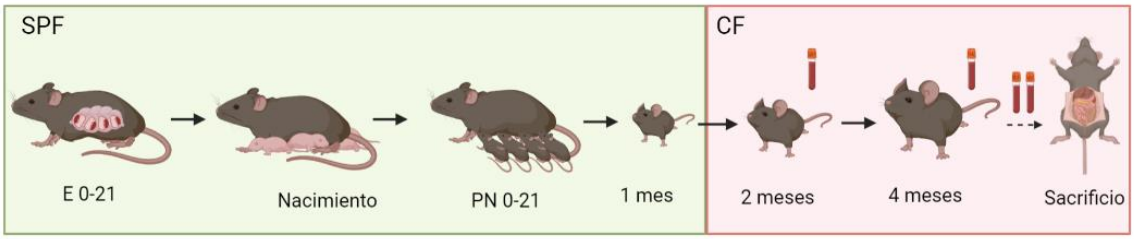
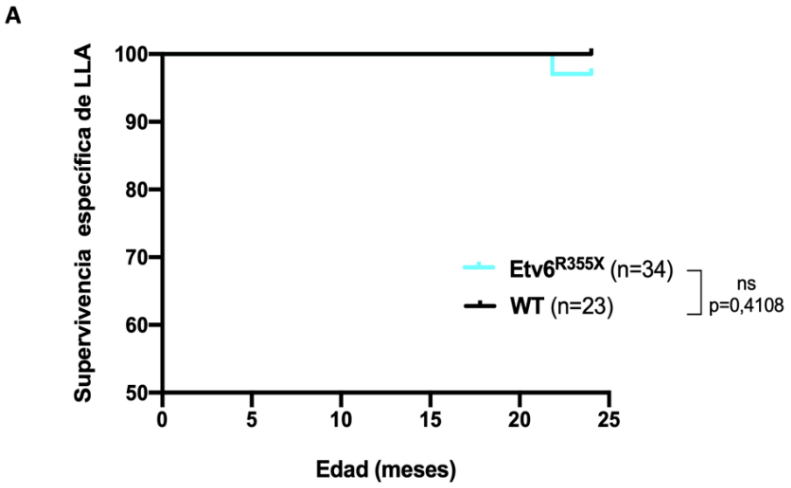



***Figura 6. Diseño experimental y seguimiento de los ratones Etv6$^{R355X}$ y WT en un animalario convencional.*** *Los ratones Etv6$^{R355X}$ y WT nacieron en el animalario SPF y aproximadamente a las 6 semanas de edad fueron trasladados a CF. Durante el estudio se realizaron análisis periódicos de la sangre periférica por citometría de flujo para determinar el estado hematopoyético de los animales y así, poder identificar la aparición de células leucémicas. Cuando los ratones alcanzaron los 22 meses de edad o cuando se detectó la LLA-B, fueron sacrificados y se procedió al análisis de los tejidos hematopoyéticos.*

Los 23 ratones WT se mantuvieron sanos durante todo el experimento y ninguno de ellos desarrolló ningún tipo de enfermedad hematológica. En el caso de los ratones *Etv6$^{R355X}$*, únicamente uno de ellos desarrolló leucemia T en el momento final del experimento (22 meses) (**Figura 7A**). De este modo, no existen diferencias estadísticamente significativas entre la incidencia de la leucemia en los ratones WT y los *Etv6$^{R355X}$*. El análisis por citometría de flujo confirmó que las células leucémicas tenían un fenotipo CD8$^{+}$ (**Figura 7B**).

Estos resultados indican que la exposición a la infección en los ratones *Etv6$^{R355X}$* con predisposición a desarrollar leucemia no es capaz de desencadenar el segundo evento que permita la conversión del clon preleucémico en leucémico.

También se estudiaron las diferentes poblaciones hematopoyéticas presentes en la sangre periférica (SP) de los ratones *Etv6$^{R355X}$* y se compararon con los de los WT. Se observó que tanto los porcentajes de células B como los de células mieloides de los ratones *Etv6$^{R355X}$* fueron superiores a los de los ratones WT en todos los puntos comparados, aunque estos datos sólo eran estadísticamente significativos en el caso de las células B (**Figura 8A y D**). En cuanto a las células T CD4$^{+}$, se observó un descenso en los porcentajes de estas células a lo largo de la vida de los ratones, tanto en los WT como en los *Etv6$^{R355X}$*, siendo siempre superiores los valores de los ratones WT (**Figura 8B**). Finalmente, los porcentajes de células T CD8$^{+}$ presentes en los WT también fueron significativamente superiores a los observados en los ratones *Etv6$^{R355X}$* en todos los puntos comparados (**Figura 8C**).



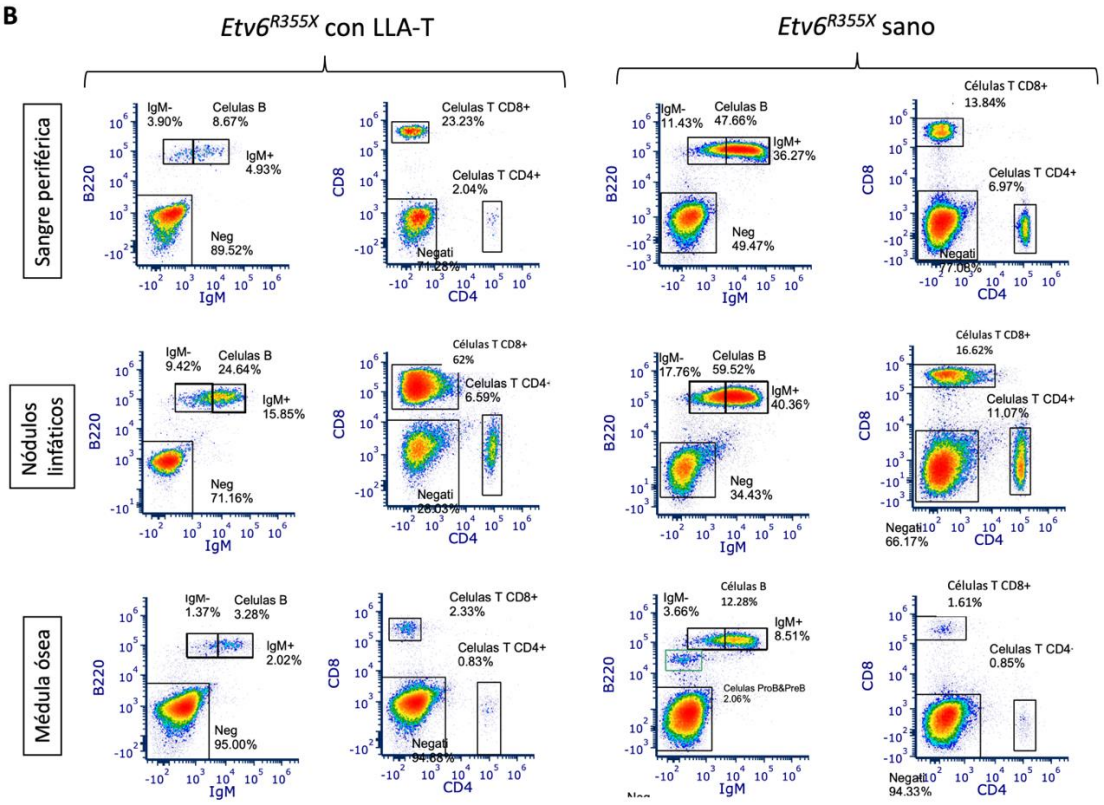




***Figura 7. Supervivencia y fenotipo de ratones Etv6$^{R355X}$ expuestos a infecciones. A)*** *Supervivencia asociada de LLA-T de ratones Etv6$^{R355X}$ (línea azul, n=34) y ratones WT (línea negra, n=23) en un animalario convencional. El p-valor fue obtenido por el análisis Log-rank (Mantel-cox): p-valor=0,4108 al comparar la supervivencia de ambos grupos de ratones.* ***B)*** *Ejemplo de un análisis de citometría de flujo de sangre periférica, nódulos linfáticos y médula ósea de un ratón Etv6$^{R355X}$ enfermo con LLA-T y otro sano de la misma edad, ambos en un animalario convencional.*

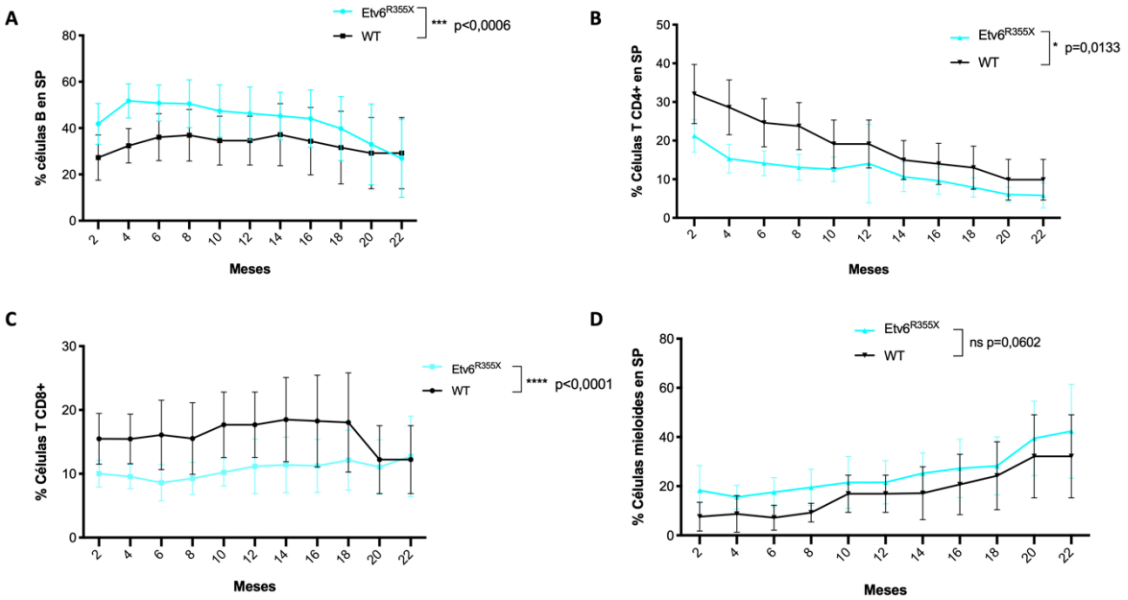



***Figura 8. Porcentajes de células B, células T y células mieloides en la sangre periférica de ratones Etv6$^{R355X}$ y ratones WT a lo largo su vida. A)*** *Se observan diferencias estadísticamente significativas entre los porcentajes de células B presentes en la sangre periférica de ratones Etv6$^{R355X}$ (línea azul) y ratones WT (línea negra) a lo largo su vida. (*** p-valor<0,0006).* ***B)*** *Se observan diferencias estadísticamente significativas entre los porcentajes de células T CD4+ presentes en la sangre periférica de ratones Etv6$^{R355X}$ y ratones WT a lo largo su vida. (* p-valor=0,0133).* ***C)*** *Se observan diferencias estadísticamente significativas entre los porcentajes de células T CD8+ presentes en la sangre periférica de ratones Etv6$^{R355X}$ y ratones WT a lo largo su vida. (**** p-valor<0,0001).* ***D)*** *No se observan diferencias estadísticamente significativas entre los porcentajes de células mieloides presentes en la sangre periférica de ratones Etv6$^{R355X}$ y ratones WT a lo largo su vida. (p-valor=0,0602).* SP: sangre periférica.

De forma paralela, se llevó a cabo un estudio de las células B preleucémicas presentes en los distintos tejidos hematopoyéticos de ratones *Etv6$^{R355X}$* jóvenes (4 meses de edad). Se observó que los porcentajes de células B en la médula ósea (MO) de los ratones *Etv6$^{R355X}$* eran inferiores a los de los ratones WT, siendo estas diferencias no estadísticamente significativas. Sin embargo, tanto en la sangre periférica, como en el bazo y en los nódulos linfáticos (NL), se encontraron valores significativamente superiores de células B en los ratones *Etv6$^{R355X}$* respecto a los WT. Además, como cabría esperar, los porcentajes de células B maduras en bazo, sangre periférica y nódulos linfáticos son superiores a los observados en la médula ósea (**Figura 9**).

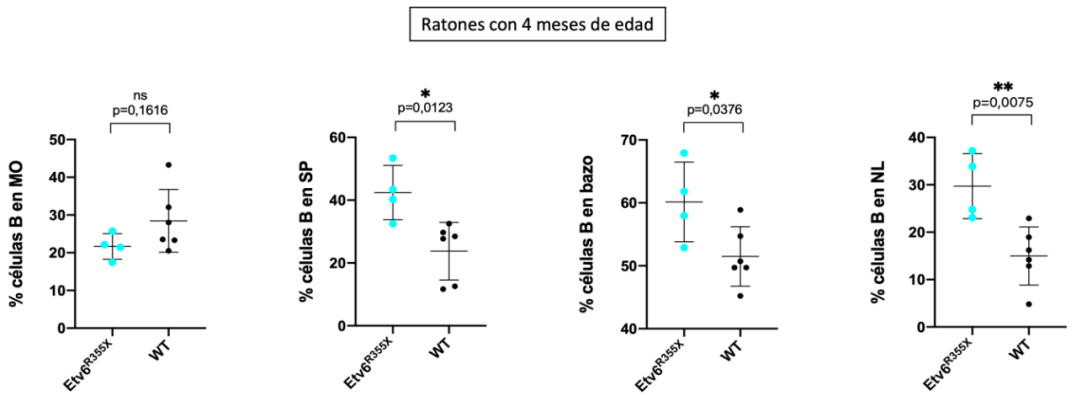



***Figura 9. Porcentajes de células B en la médula ósea, sangre periférica, bazo y nódulos linfáticos de ratones Etv6$^{R355X}$ y WT de 4 meses de edad.*** *No se observaron diferencias en la población B de la médula ósea (MO) de ratones Etv6$^{R355X}$ (n=4) en comparación con ratones WT (n=6). Si hay diferencias estadísticamente significativas entre los porcentajes de células B presentes en la sangre periférica (SP) de ratones Etv6$^{R355X}$ en comparación con ratones WT (* p-valor=0,0123). Si hay diferencias estadísticamente significativas entre los porcentajes de células B presentes en el bazo de ratones Etv6$^{R355X}$ respecto a los WT (* p-valor=0,0376). También hay diferencias estadísticamente significativas entre los porcentajes de células B presentes en los nódulos linfáticos (NL) de ratones Etv6$^{R355X}$ en comparación con los WT (** p-valor=0,0075). Todos los ratones fueron mantenidos en un animalario convencional. Los p-valores corresponden a los obtenidos por un análisis t-test no pareado.*

Finalmente, también se llevó a cabo un análisis donde se compararon las poblaciones de células B de la médula ósea de ratones *Etv6$^{R355X}$* jóvenes (4 meses de edad) y ratones mayores (22 meses de edad) para ver cómo variaban los porcentajes de estas poblaciones con el paso del tiempo. Como cabría esperar, se observó una disminución de las células B totales en la médula ósea como consecuencia del envejecimiento de los animales (**Figura 10A**). De igual manera, cuando se separó el análisis atendiendo a los diferentes compartimentos de células B presentes en la médula ósea en función de la expresión de los marcadores B220 e IgM, se observaron resultados parecidos. Los ratones *Etv6$^{R355X}$* de 4 meses de edad presentaban un mayor porcentaje de células preB y proB (B220$^{bajo}$ IgM$^{-}$) y de células inmaduras (B220$^{bajo}$ IgM$^{+}$) que los ratones de más edad (**Figura 10A**). Sin embargo, los ratones jóvenes mostraban menor porcentaje de células recirculantes (B220$^{+}$ IgM$^{+/-}$) que los ratones de 22 meses, aunque estas diferencias no eran estadísticamente significativas (**Figura 10A**).

Estos datos también se reflejaban en los análisis de citometría de flujo de médula ósea de ratones *Etv6$^{R355X}$* de 4 meses de edad y 22 meses de edad. Según lo esperado, el ratón joven presentaba un mayor porcentaje de células B inmaduras y células preB y proB en la médula ósea, pero es el ratón de 22 meses el que muestra porcentaje mayor de células B maduras recirculantes (**Figura 10B**).

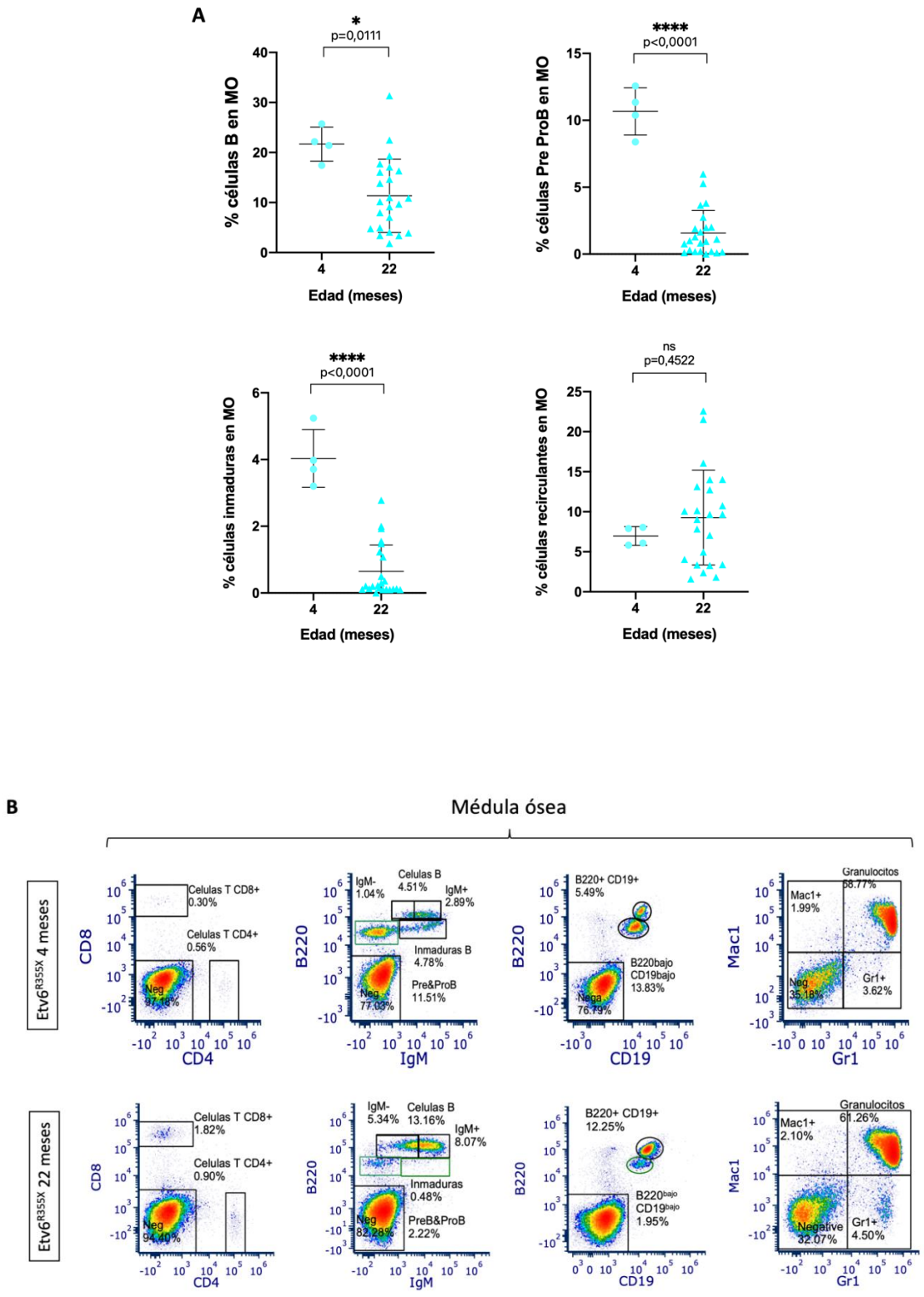



***Figura 10. Análisis comparativo de la médula ósea entre ratones Etv6$^{R355X}$ de 4 y 22 meses de edad. A)*** *Análisis estadístico de los porcentajes de células B totales, de células preB y proB, inmaduras y recirculantes presentes en la médula ósea (MO) de ratones Etv6$^{R355X}$ a los 4 (n=4) y 22 (n=23) meses de edad. Los p-valores corresponden a los obtenidos por un análisis t-test no pareado.* ***B)*** *Ejemplo de análisis de citometría de flujo entre un ratón Etv6$^{R355X}$ de 4 meses y otro de 22 meses en el que se muestran las poblaciones de células T, células B y células mieloides presentes en la MO de estos ratones.*

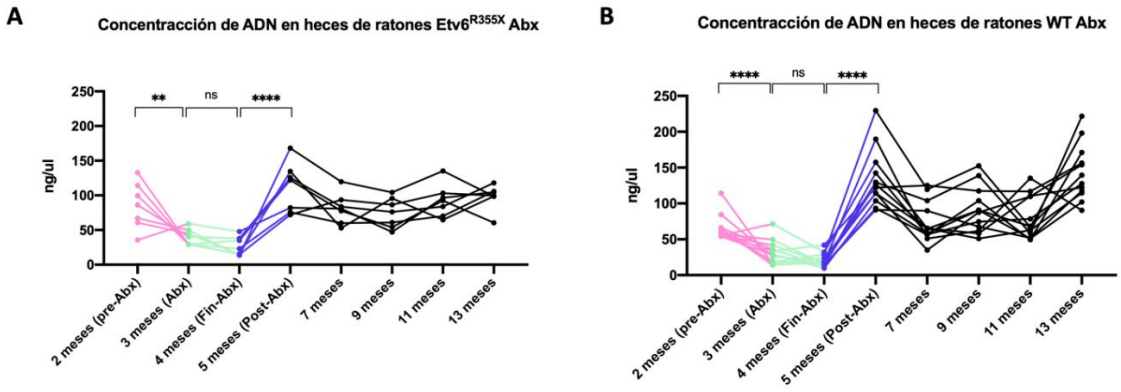
## 1.2. La alteración del microbioma intestinal no promueve el desarrollo de la leucemia en ratones *Etv6*$^{R355X}$ genéticamente predispuestos.

Actualmente, se sabe que tanto los ratones *Pax5*$^{+/-}$ como los *Sca1-ETV6-RUNX1* presentan microbiomas intestinales específicos de cada genotipo directamente relacionados con la alteración genética que portan. Además, en el caso de los ratones *Pax5*$^{+/-}$, su alteración mediante el tratamiento con antibióticos es suficiente para inducir la LLA-B incluso en ausencia de estímulos infecciosos[100].

Estudios previos han publicado que las LLA-B diploides asociadas a *ETV6* presentan la deleción de *PAX5* con una frecuencia del 30%[80]. Teniendo esto en cuenta, se decidió estudiar si el estrés inmunológico inducido con el uso de los antibióticos en condiciones SPF podría conducir al desarrollo de la leucemia en ratones *Etv6*$^{R355X}$ genéticamente predispuestos. Para este experimento se generaron 32 ratones *Etv6*$^{R355X}$ y 12 ratones WT, ambas cohortes fueron tratadas con antibióticos. También se generaron 35 ratones *Etv6*$^{R355X}$ que conformaron el grupo control al no recibir el tratamiento antibiótico. Sólo se utilizaron hembras en el estudio para evitar el sesgo por razón de sexo, ya que el microbioma intestinal es diferente entre machos y hembras. Todos los ratones se mantuvieron en un animalario SPF durante todo el experimento. En torno a las 8 semanas de edad, los 32 ratones *Etv6*$^{R355X}$ y los 12 ratones WT fueron tratados con un cóctel de antibióticos (ampicilina, vancomicina, ciprofloxacina, imipenem y metronidazol) disuelto en el agua de los biberones durante 8 semanas, con el objetivo de eliminar por completo los microorganismos que conformaban su microbiota intestinal. Se recogieron muestras fecales (de las 3 cohortes de ratones) antes de empezar el tratamiento, a las 4 semanas de tratamiento, a las 8 semanas de tratamiento (coincidiendo con el fin de éste), 4 semanas tras la finalización del tratamiento y después, cada 2 meses hasta que los animales mostraron signos de enfermedad o hasta el punto final del experimento (22 meses). Además, periódicamente se llevaron a cabo análisis de sangre periférica mediante citometría de flujo para comprobar cómo afectaba el tratamiento con los antibióticos a la composición hematopoyética y así poder detectar la presencia de células leucémicas en sangre (**Figura 11**).

***Figura 11. Diseño experimental del tratamiento con antibióticos y seguimiento de los ratones.*** *Los ratones Etv6$^{R355X}$ y WT nacieron y se mantuvieron en el animalario SPF toda su vida. A los 2 meses de edad si inició el tratamiento con el cóctel de antibióticos con una duración de 8 semanas. Durante este tiempo se recogieron muestras fecales (al inicio del tratamiento, a las 4 semanas, al acabar el tratamiento, un mes después de acabar el tratamiento y después cada 2 meses). También se realizaron análisis periódicos mediante citometría de flujo de la sangre periférica cada 2 meses para hacer un seguimiento del estado hematopoyético de los animales y poder así encontrar las células leucémicas en circulación. Bien cuando se detectaron casos de leucemia o bien cuando los ratones alcanzaron los 22 meses de edad, fueron sacrificados para hacer un análisis de los distintos órganos hematopoyéticos por citometría de flujo.* SP: sangre periférica.

El tratamiento con antibióticos fue eficaz en cuanto a la eliminación del microbioma intestinal. La ausencia de éste se determinó midiendo la concentración del ADN presente en las heces de los animales (**Figura 12A-B**). Para el caso de los ratones *Etv6$^{R355X}$* y WT tratados con antibióticos se observó que tras 4 semanas de tratamiento ya había una reducción significativa en la concentración de ADN en las muestras fecales; y tras los dos meses con los antibióticos la concentración se redujo hasta valores próximos a 0. Además, un mes después de la finalización del tratamiento con los antibióticos los valores se restablecieron (**Figura 12A-B**). Sin embargo, en el grupo de ratones control *Etv6$^{R355X}$*, a los que no se les administraron los antibióticos, las concentraciones de ADN bacteriano de las muestras fecales se mantienen sin variaciones significativas a lo largo del tiempo (**Figura 12C**).



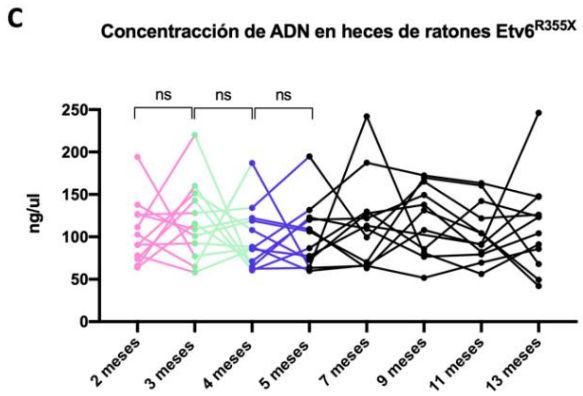
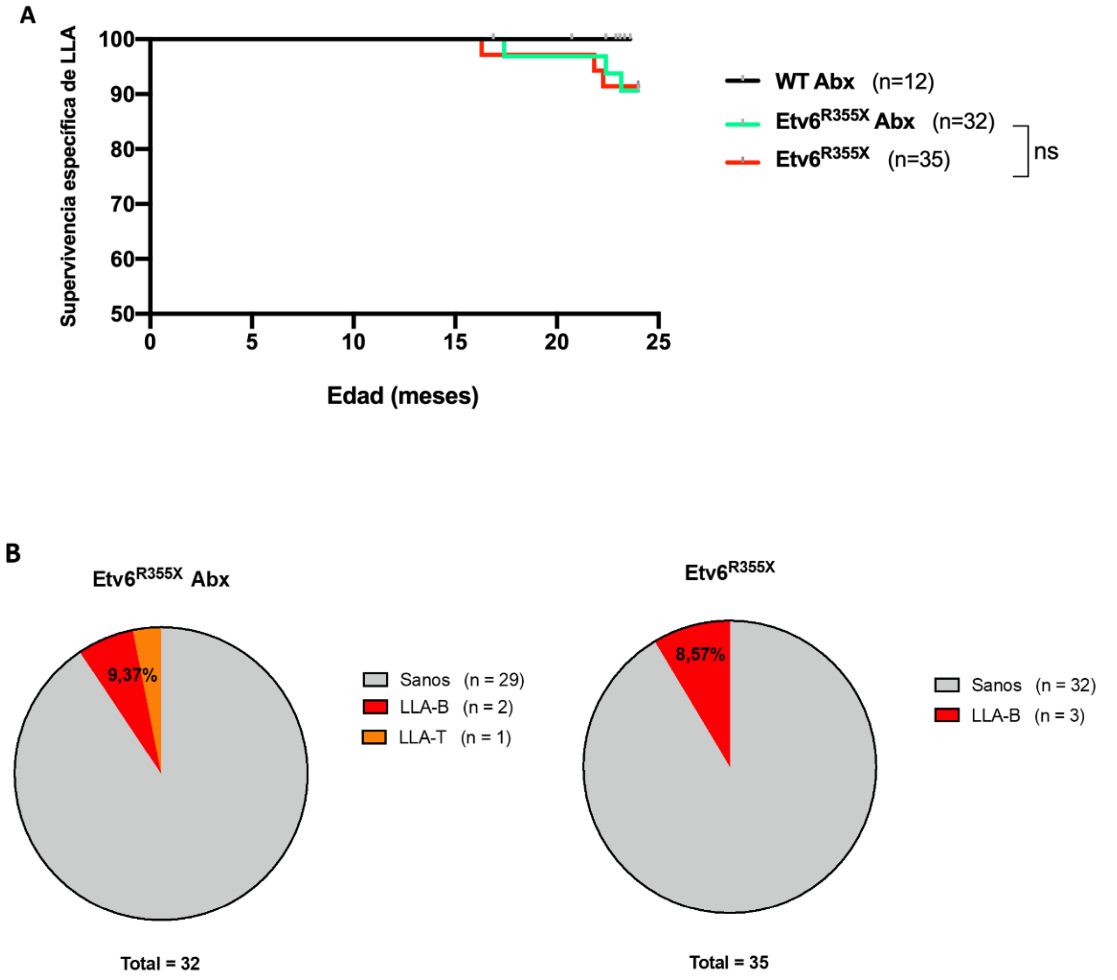


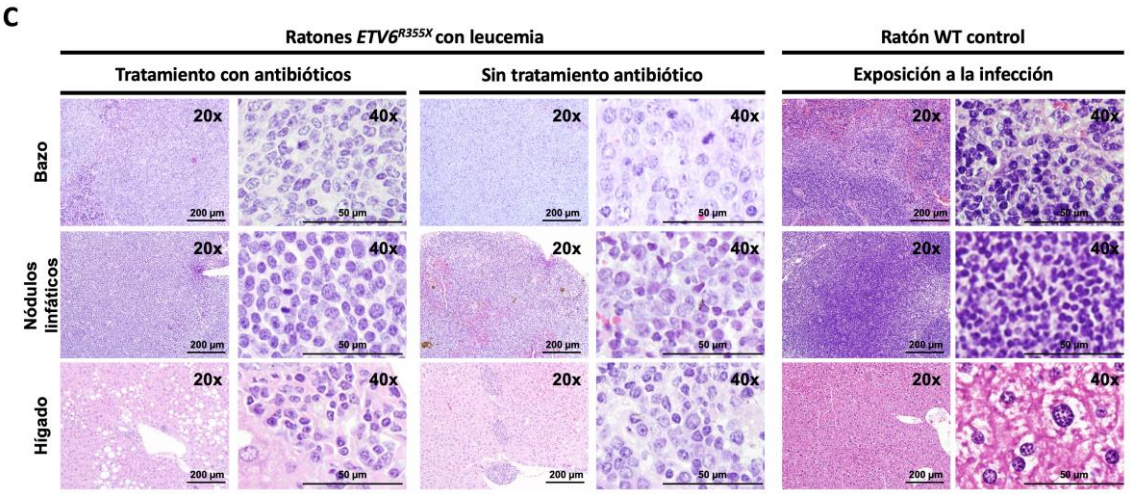

***Figura 12. Gráficas representativas de la concentración de ADN en las muestras fecales recogidas a lo largo del experimento en ratones Etv6$^{R355X}$ y WT. A)*** *Representación de la concentración de ADN de las muestras fecales de los ratones Etv6$^{R355X}$ tratados con antibióticos a lo largo del tiempo (los meses indican la edad de los ratones). Al finalizar el tratamiento con antibióticos (4 meses de edad) la concentración de ADN en las heces era prácticamente 0. Las concentraciones de ADN bacteriano en las muestras fecales se reestablecieron aproximadamente 1 mes tras finalizar el tratamiento (5 meses de edad). (** p-valor=0,0057; **** p-valor<0,0001)* ***B)*** *Concentración de ADN en las muestras fecales de los ratones WT tratados con antibióticos a lo largo del tiempo. Al finalizar el tratamiento con antibióticos (4 meses de edad) la concentración de ADN en las heces era prácticamente 0. Las concentraciones de ADN bacteriano en las muestras fecales se reestablecieron aproximadamente 1 mes tras finalizar el tratamiento. (**** p-valor<0,0001).* ***C)*** *Representación de la concentración de ADN de las muestras fecales de los ratones Etv6$^{R355X}$ no tratados con antibióticos. Los valores de las concentraciones se mantienen sin variaciones significativas a lo largo del tiempo. Cada punto hace referencia a la media de las concentraciones de ADN de los ratones mantenidos en la misma jaula. Abx: tratamiento con antibióticos.*

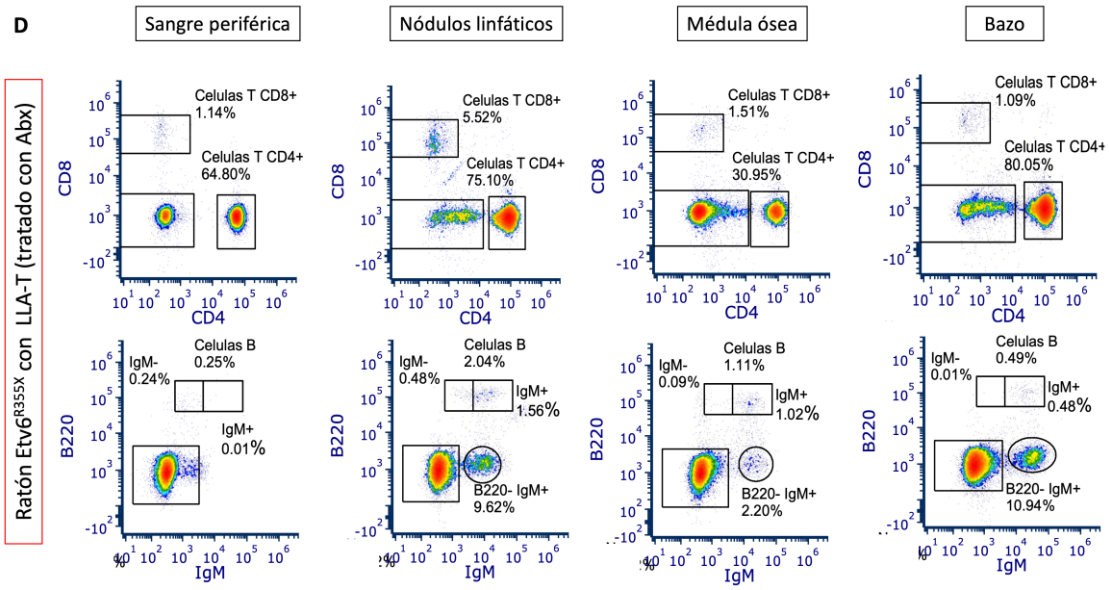
Resultados previos habían demostrado que en el caso de los ratones *Pax5$^{+/-}$* genéticamente predispuestos, la depleción bacteriana debida al tratamiento con antibióticos desencadenaba la LLA-B, incluso en condiciones SPF, es decir, en ausencia de estrés inmune debido a la exposición a infecciones comunes[100]. Para corroborar si este hecho también se daba en los ratones *Etv6$^{R355X}$*, se decidió explorar si la alteración del microbioma intestinal de estos ratones también inducía la leucemia. Los resultados mostraron que tanto la incidencia como el tiempo transcurrido hasta la aparición de la enfermedad no se vieron alterados por el tratamiento con antibióticos al compararlos con los ratones que no siguieron el tratamiento (**Figura 13A**). En el caso de los ratones *Etv6$^{R355X}$* tratados con antibióticos, dos de ellos desarrollaron LLA-B y uno LLA-T, siendo la latencia de la enfermedad de entre 17 y 23 meses de edad. En el grupo de los ratones *Etv6$^{R355X}$* que no vieron alterado su microbioma intestinal, hubo 3 casos de LLA-B que se detectaron entre los 16 y los 22 meses de edad (**Figura 13A-B**). En ambos grupos, las leucemias se diagnosticaron por la aparición de células tumorales en la sangre periférica, la médula ósea, el bazo o los nódulos linfáticos, las cuales llegaron a infiltrar otros tejidos no hematopoyéticos como el hígado (**Figura 13C**). Desde el punto de vista fenotípico, la

LLA-T presentaba una población tumoral B220$^{-}$ IgM$^{+}$ CD4$^{+}$ (**Figura 13D**) y las LLA-B eran B220$^{bajo}$ IgM$^{-}$ CD19$^{bajo/+}$ Mac1$^{+}$ (**Figura 13E**). Por lo tanto, este estudio indica que no existen diferencias con relación a la incidencia de la leucemia entre los ratones *Etv6*$^{R355X}$ tratados y no tratados con antibióticos, reflejando que, en este caso, un estrés inmunológico como es el uso de antibióticos no es un factor diferencial a la hora de disparar la leucemia en ratones con una misma predisposición genética.

Para confirmar que las leucemias de los ratones *Etv6*$^{R355X}$ realmente pueden ser catalogadas como tal, es necesario llevar a cabo la identificación de los eventos secundarios que se adquieren somáticamente y que conducen al desarrollo de la LLA. Para ello se requiere secuenciar el genoma completo de las muestras de ADN tumoral y compararlos con sus correspondientes muestras de ADN control obtenidas de la línea germinal de la cola de dichos ratones. Este es un proceso largo por lo que actualmente no se dispone de los resultados de la secuenciación, pero estamos a la espera de ellos para confirmar así la aparición de la leucemia en los ratones *Etv6*$^{R355X}$.




C
Ratones ETV6R355X con leucemia
Ratón WT control
Tratamiento con antibióticos
Sin tratamiento antibiótico
Exposición a la infección
Bazo
Nódulos linfáticos
Hígado
20x
40x
200 µm
50 µm


D
Sangre periférica
Nódulos linfáticos
Médula ósea
Bazo
Ratón Etv6R355X con LLA-T (tratado con Abx)
CD8
CD4
B220
IgM
Celulas T CD8+ 1.14%
Celulas T CD4+ 64.80%
Celulas B 0.25%
IgM- 0.24%
IgM+ 0.01%
Celulas T CD8+ 5.52%
Celulas T CD4+ 75.10%
Celulas B 2.04%
IgM- 0.48%
IgM+ 1.56%
B220- IgM+ 9.62%
Celulas T CD8+ 1.51%
Celulas T CD4+ 30.95%
Celulas B 1.11%
IgM- 0.09%
IgM+ 1.02%
B220- IgM+ 2.20%
Celulas T CD8+ 1.09%
Celulas T CD4+ 80.05%
Celulas B 0.49%
IgM- 0.01%
IgM+ 0.48%
B220- IgM+ 10.94%

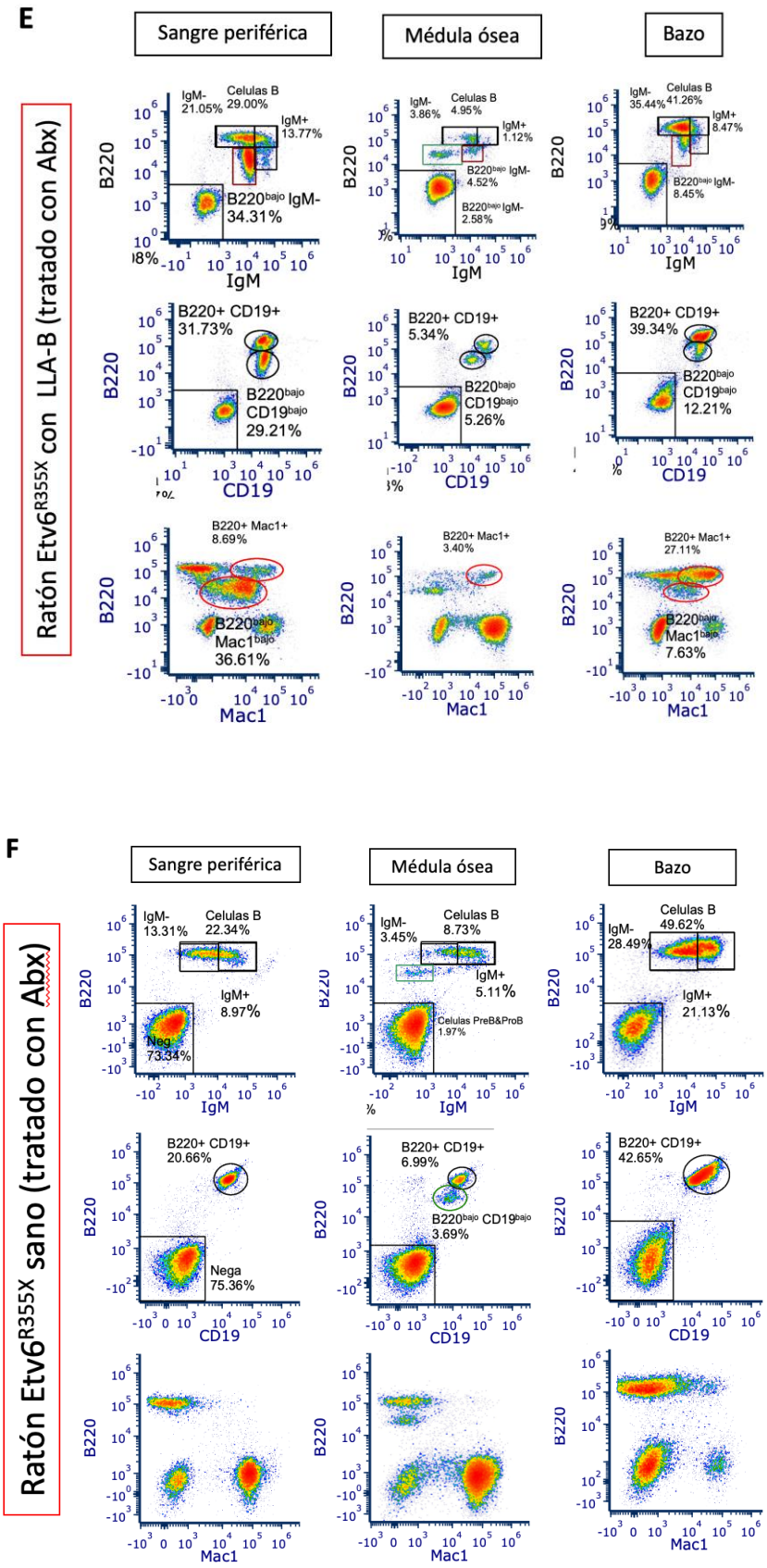



***Figura 13. Curva de supervivencia específica, incidencia de leucemia y fenotipado de ratones Etv6$^{R355X}$ con y sin tratamiento con antibióticos. A)*** *Curva de supervivencia específica de LLA en ratones Etv6$^{R355X}$ tratados (n=32) o no tratados (n=35) con antibióticos (p-value=0,9347) y WT tratados (n=12). Ninguno de ellos fue expuesto a infecciones comunes.* ***B)*** *Incidencia de leucemia en ratones Etv6$^{R355X}$ con y sin privación de la microbiota intestinal. Todos los*

*ratones fueron alojados en condiciones SPF. No se encontraron diferencias significativas mediante la prueba exacta de Fisher. El 9,37% (3/32) de los Etv6$^{R355X}$ privados del microbioma desarrollaron LLA (2 LLA-B y 1 LLA-T) y el 8,57% (3/35) de los Etv6$^{R355X}$ no tratados desarrollaron LLA-B.* ***C)*** *Tinción con hematoxilina y eosina de distintos tejidos de ratones leucémicos Etv6$^{R355X}$ tratados con y sin antibióticos y comparados con los de un ratón WT. Se muestran imágenes de las histologías de bazo, nódulos linfáticos e hígado. Se aprecia la pérdida de la arquitectura normal de los tejidos en los ratones con LLA-B en comparación con los ratones sanos. En cada caso se indica la ampliación y la escala.* ***D)*** *Ejemplo representativo de análisis por citometría de flujo de sangre periférica (primera columna), nódulos linfáticos (segunda columna), médula ósea (tercera columna) y bazo (cuarta columna) de un ratón Etv6$^{R355X}$ tratado con antibióticos y que presenta LLA-T.* ***E)*** *Ejemplo de análisis por citometría de flujo de sangre periférica (primera columna), médula ósea (segunda columna) y bazo (tercera columna) de un ratón Etv6$^{R355X}$ tratado con antibióticos y que presenta LLA-B.* ***F)*** *Ejemplo de análisis por citometría de flujo de sangre periférica (primera columna), médula ósea (segunda columna) y bazo (tercera columna) de un ratón Etv6$^{R355X}$ sano tratado con antibióticos. En los ejes de las figuras se indican los marcadores celulares analizados. Todos los ratones fueron mantenidos en condiciones SPF toda su vida. Abx: tratados con un cóctel de antibióticos durante 8 semanas.*

Ninguno de los ratones WT tratados con el cóctel de antibióticos durante las 8 semanas desarrollaron LLA-B, pero 4 de ellos (4 de 12) presentaron otros tumores cuando tenían 22 meses de edad, los cuales fueron detectados en el momento del sacrificio del ratón por el fin del experimento. Los análisis de inmunofenotipado por citometría de flujo confirmaron que los tumores eran neoplasias de células B maduras (linfomas), y que por lo tanto eran distintas a las leucemias observadas en los ratones *Etv6$^{R355X}$* (**Figura 14**). Estos resultados confirman que la eliminación del microbioma intestinal no tiene un papel disparador de la leucemia en ratones sin predisposición genética a esta enfermedad.

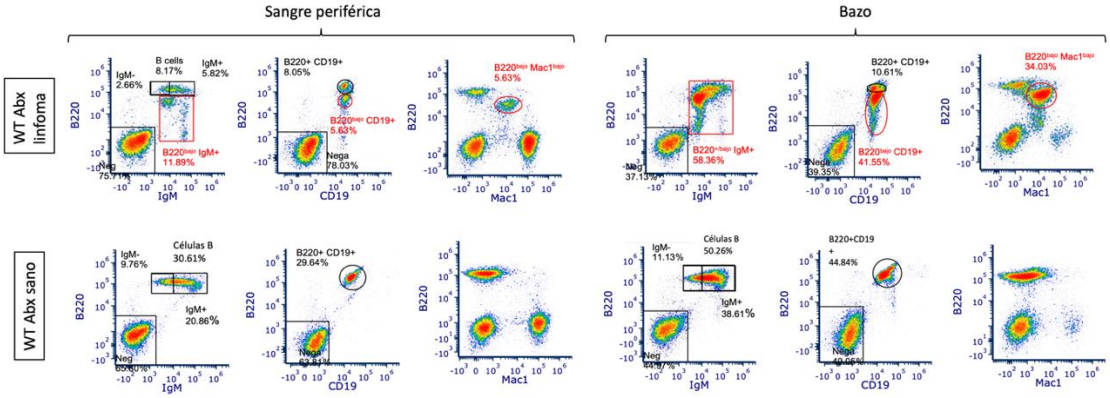



***Figura 14. Análisis de la enfermedad en ratones WT estabulados en SPF y que han sufrido la eliminación del microbioma mediante el uso de antibióticos.*** *Ejemplo representativo del análisis por citometría de flujo de sangre periférica y bazo de un ratón WT tratado con antibióticos que presenta linfoma de células B (primera fila) y otro ratón WT sano de la misma edad (segunda fila). En los ejes de las figuras se indica el marcador celular utilizado en cada caso. Todos los ratones fueron mantenidos en SPF toda su vida. Abx: tratamiento con antibióticos.*

Cuando se analizó el efecto del tratamiento con antibióticos en la composición hematopoyética de los ratones, se observó un aumento significativo a los 6 y a los 8 meses de edad en el número de células B maduras en sangre periférica en los ratones *Etv6*$^{R355X}$ que habían sido tratados con antibióticos (**Figura 15**). Sin embargo, en cuanto a las células T CD8$^{+}$ y a las células mieloides, se detectó una disminución de ambas poblaciones, también a los 6 y 8 meses de edad, en los ratones *Etv6*$^{R355X}$ tratados con antibióticos respecto de los que no habían sufrido privación del microbioma intestinal (**Figura 15**). El hecho de que los antibióticos fueron administrados hasta los 4 meses de edad y que las variaciones estadísticamente significativas de las distintas poblaciones hematopoyéticas entre ratones tratados y no tratados se observaron únicamente en los meses próximos a la finalización del tratamiento, puede sugerir que estas diferencias se deban al efecto de los antibióticos. De esta manera, a medida que pasa el tiempo, su efecto se va diluyendo y los porcentajes de las poblaciones celulares presentes en sangre se van restaurando hasta asemejarse a los de los ratones que no han seguido el tratamiento con antibióticos.

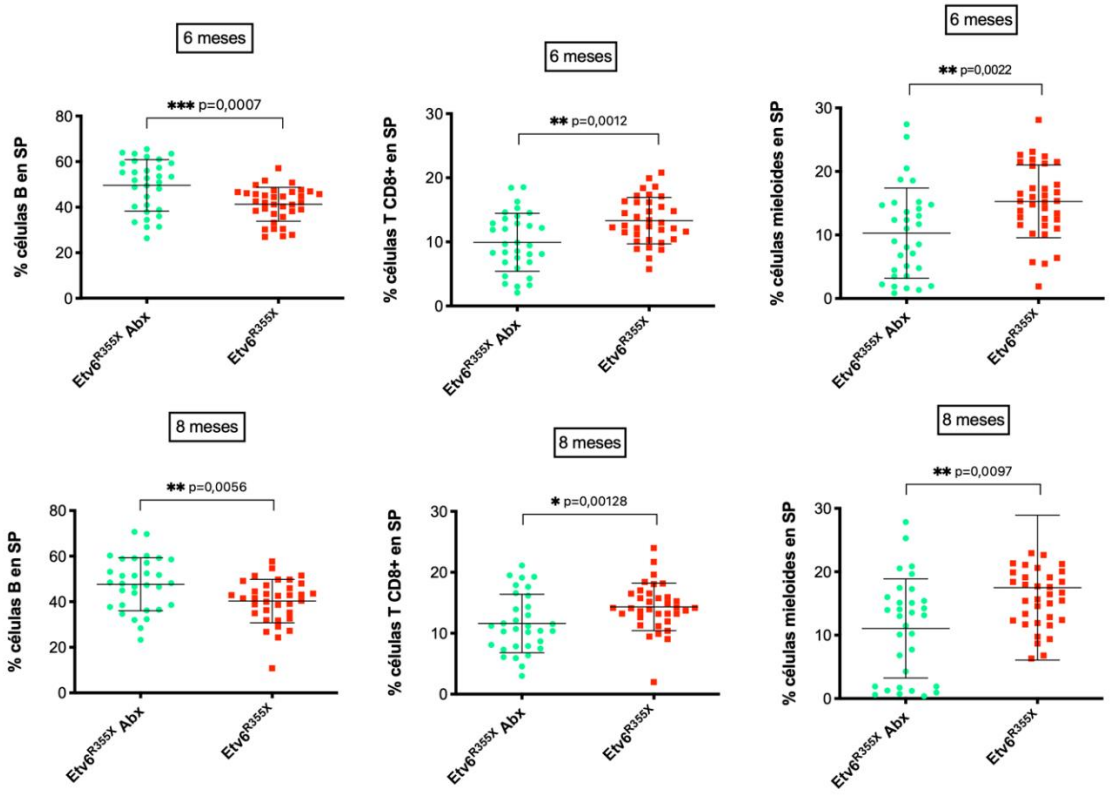


***Figura 15. Análisis representativos estadísticos de los cambios producidos por el tratamiento con antibióticos (Abx) en las poblaciones hematopoyéticas de sangre periférica en los ratones Etv6$^{R355X}$.*** *Análisis estadístico de las poblaciones hematopoyéticas de sangre periférica (SP) en las que se observaron cambios estadísticamente significativos debido al tratamiento con antibióticos. Se observaron diferencias en la población B (primera columna) de sangre periférica en ratones Etv6$^{R355X}$ tratados con antibióticos (n=32) en comparación con ratones del mismo genotipo sin tratamiento con antibióticos (n=35). Se observó una disminución significativa en el número de células B en los ratones Etv6$^{R355X}$ no tratados con antibióticos comparado con ratones Etv6$^{R355X}$ tratados a los 6 meses (primera fila; p-valor=0,0007) y a los 8 meses (segunda fila; p-valor=0,0064). Se observó también una disminución de células T CD8+ (segunda columna) debido al tratamiento con antibióticos (p-valor=0,0012 a los 6 meses y p-valor=0,00128 a los 8 meses de edad). Se observó una disminución significativa en el número de células mieloides en los ratones Etv6$^{R355X}$ tratados con antibióticos comparado con ratones Etv6$^{R355X}$ no tratados tanto a los 6 meses (primera fila; p-valor=0,0022) como a los 8 meses (segunda fila; p-valor=0,0097). Los p-valores corresponden a los obtenidos por test t no pareado.*

Se quisieron también analizar las posibles diferencias en cuanto a la composición de células B presentes en la médula ósea de los ratones *Etv6$^{R355X}$* con y sin privación del microbioma intestinal. Se observó que el porcentaje de células B totales en la médula ósea de los ratones *Etv6$^{R355X}$* tratados con antibióticos era menor que el que presentaban los animales del mismo genotipo no tratados y menor también que el de los ratones WT tratados con antibióticos, pero estas diferencias no eran estadísticamente significativas (**Figura 16A**). De igual manera, cuando se separa el análisis atendiendo a los diferentes compartimentos de células B presentes en la médula ósea en función de la expresión de los marcadores B220 e IgM, se observaron resultados parecidos. Los ratones *Etv6$^{R355X}$* con privación del microbioma intestinal presentan un menor porcentaje de células preB y proB (B220$^{bajo}$ IgM$^{-}$), de células inmaduras (B220$^{bajo}$ IgM$^{+}$) y de células recirculantes (B220$^{+}$ IgM$^{+/-}$) que los ratones *Etv6$^{R355X}$* no tratados y que los WT con tratamiento antibiótico. En ningún caso estas diferencias son estadísticamente significativas **(Figura 16B**).

Por lo tanto, ni la diferencia entre genotipos (*Etv6$^{R355X}$* vs WT) ni la presencia o ausencia de un estrés inmunológico (tratamiento con antibióticos) dentro del mismo genotipo, afecta en este caso a las poblaciones de células B presentes en la médula ósea. Este análisis se realizó con médulas óseas de ratones sanos que llegaron al punto final de fin de experimento (22 meses), de manera que el hecho de no observar diferencias significativas también puede deberse en parte a que el efecto transitorio de los antibióticos ya no esté presente.

Finalmente, también se llevó a cabo un análisis donde se comparó la médula ósea de ratones jóvenes (4 meses de edad) y ratones mayores (22 meses) *Etv6$^{R355X}$* para ver

si la edad de los animales podía tener un efecto sobre las poblaciones de células B presentes en este tejido hematológico. No se observaron diferencias significativas ni entre las células B totales ni entre las diferentes poblaciones de células inmaduras presentes en la médula ósea de los ratones *Etv6*$^{R355X}$ a los dos puntos de edad (**Figura 16C**).

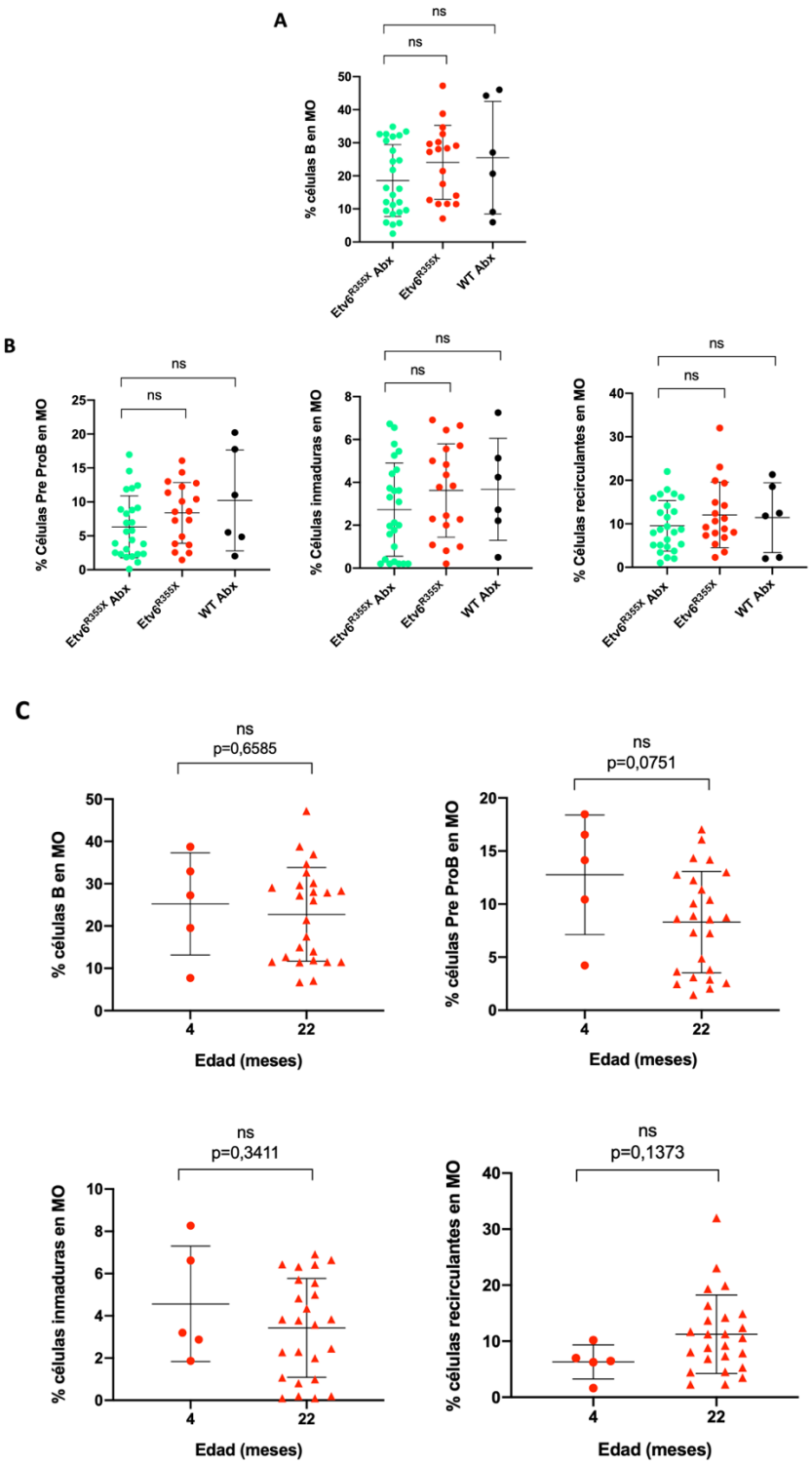
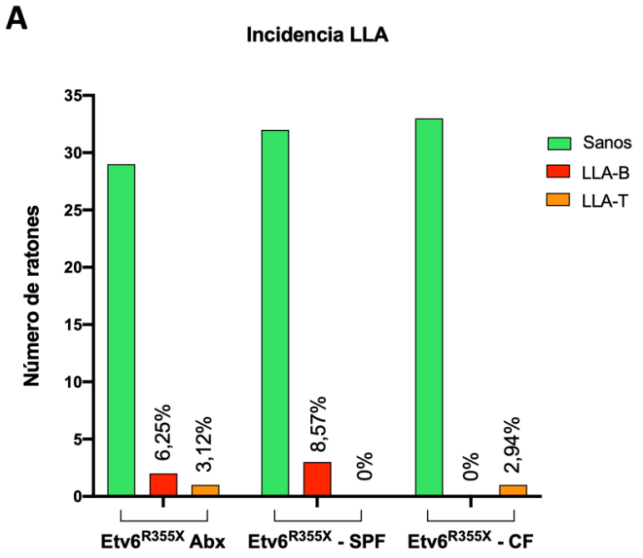


***Figura 16. Porcentajes de células B en la médula ósea de ratones Etv6$^{R355X}$ y WT tratados y no tratados con antibióticos. A)*** *Porcentajes de células B totales en médula ósea (MO) en ratones Etv6$^{R355X}$ con tratamiento (n=25), sin tratamiento (n=18) y WT con tratamiento (n=6), a los 22 meses de edad.* ***B)*** *Porcentajes de células preB y proB, inmaduras y recirculantes en MO en ratones con tratamiento (n=25), sin tratamiento (n=18) y WT con tratamiento (n=6), a los 22 meses de edad.* ***C)*** *Comparación entre los porcentajes de células B totales, de células preB y proB, inmaduras y recirculantes presentes en la MO de ratones Etv6$^{R355X}$ sin tratar con antibióticos, a los 4 (n=5) y 22 (n=25) meses de edad. Los p-valores corresponden a los obtenidos por análisis t-test no pareado.*

Una vez estudiado el papel de distintos estreses inmunológicos, como la exposición a infecciones comunes y el tratamiento con antibióticos, en ratones *Etv6$^{R355X}$* predispuestos a desarrollar LLA-B, se concluyó que no se dan cambios significativos en cuanto a la incidencia de leucemia respecto a la de los grupos de ratones control. De manera que, en el caso de los ratones *Etv6$^{R355X}$* expuestos a patógenos, únicamente un ratón desarrolló LLA-T y por lo tanto no existen diferencias significativas en relación con la incidencia entre estos ratones y los WT también expuestos a infecciones. Respecto al inmuno-estrés causado por la depleción bacteriana, la incidencia de la enfermedad tampoco se vio alterada por el tratamiento con antibióticos en comparación con los ratones *Etv6$^{R355X}$* que no siguieron dicho tratamiento (**Figura 17A**). Del mismo modo, los distintos porcentajes de supervivencia son muy similares entre los ratones *Etv6$^{R355X}$* de CF y los ratones *Etv6$^{R355X}$* de SPF, con y sin tratamiento con antibióticos, no habiendo diferencias estadísticamente significativas entre ellos (**Figura 17B**). Es por ello por lo que se puede concluir que el estrés inmunológico no parece tener un papel promotor de la leucemia vinculada a la predisposición genética del gen *ETV6*.



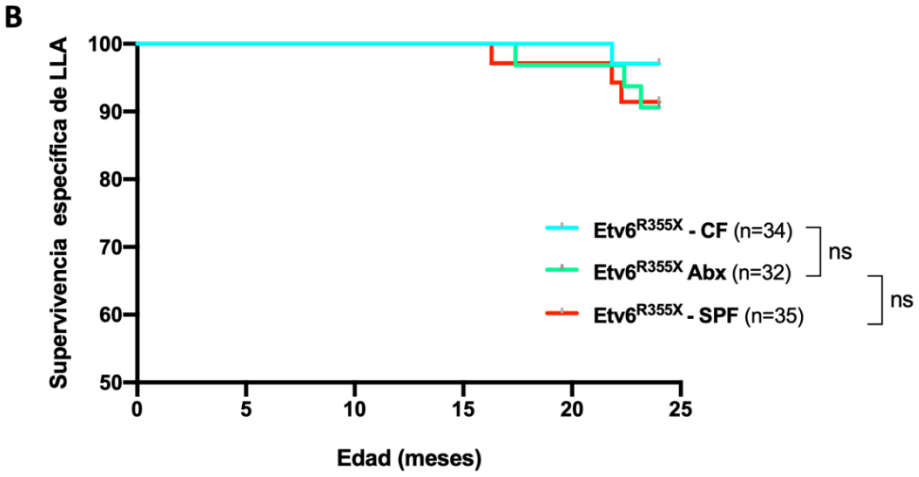



***Figura 17. Comparación de la incidencia y supervivencia de ratones Etv6$^{R355X}$ expuestos a distintos estreses inmunológicos. A)*** *Incidencia de LLA-B (6,25%) y LLA-T (3,12%) en ratones Etv6$^{R355X}$ con tratamiento con antibióticos. Incidencia de LLA-B (8,57%) que presentan los ratones Etv6$^{R355X}$ que no han tenido privación del microbioma intestinal. Incidencia de LLA-T (3,12%) de los ratones Etv6$^{R355X}$ expuestos a infecciones comunes.* ***B)*** *Curva de supervivencia específica de LLA en ratones Etv6$^{R355X}$ tratados (línea verde, n=32), no tratados (línea roja, n=35) con antibióticos y expuestos a patógenos (línea azul, n=34). El p-valor fue obtenido por el análisis Log-rank (Mantel-cox): p-valor=0,9347 al comparar la supervivencia de Etv6$^{R355X}$ Abx vs Etv6$^{R355X}$ (SPF); p-valor=0,2808 al comparar la supervivencia de Etv6$^{R355X}$ Abx vs Etv6$^{R355X}$ (CF). Abx: tratamiento con antibióticos.*

## 2. El estrés inmune suprime la señalización de la inmunidad innata en las células B precursoras preleucémicas para causar leucemia.

La LLA-B es considerada un trastorno sanguíneo clonal cuyas fases más tempranas tienen lugar durante el desarrollo embrionario. Las células preleucémicas, portadoras de alteraciones que afectan a genes clave para el desarrollo hematopoyético tempano, son compatibles con una hematopoyesis normal en la mayoría de las ocasiones, sin embargo, a veces pueden transformarse y dar lugar a la aparición de la LLA-B gracias a la adquisición de mutaciones somáticas adicionales[10,25,26,86,101]. Estudios recientes han demostrado que el estrés inmune causado por la exposición ambiental a infecciones comunes proporciona una ventaja de selección a las células preleucémicas que mediante un mecanismo de transformación dan lugar a la leucemia. Sin embargo, todavía se desconoce cómo este estrés desencadena la transformación[25,26,86], aunque

recientemente se ha empezado a estudiar cómo las células preleucémicas responden a los estímulos infecciosos.

Este trabajo trata de determinar cómo las células preleucémicas *Pax5*$^{+/-}$ y *ETV6-RUNX1* pasan a un estado leucémico durante el estrés inmunitario, examinando para ello ratones transgénicos *Pax5*$^{+/-}$ y *ETV6-RUNX1* susceptibles a desarrollar leucemia. Se ha visto que estos modelos murinos reproducen fielmente el desarrollo de la LLA-B que se da en humanos.

Además, se sabe que la presencia de las mutaciones *Pax5*$^{+/-}$ o *ETV6-RUNX1* modifica de forma única el microbioma intestinal en comparación con sus homólogos WT. Curiosamente, este microbioma intestinal específico de los ratones *Pax5*$^{+/-}$ les protege contra el desarrollo de la leucemia, y su alteración mediante el tratamiento con antibióticos es suficiente para inducir la LLA-B, incluso en ausencia de estímulos infecciosos[32]. Existe también una alianza entre el sistema inmune y el microbioma intestinal, que cuando funciona de forma óptima, permite la inducción de respuestas protectoras frente a patógenos y el mantenimiento de vías reguladoras implicadas en el mantenimiento de la tolerancia a antígenos inocuos[102]. Sin embargo, se ha demostrado que la alteración bacteriana provoca una supresión inmunitaria tanto innata como adaptativa[103].

La inmunidad innata constituye la primera línea de defensa contra los microorganismos invasores y su importancia en la formación de tumores también se pone de manifiesto por el creciente número de cánceres (por ejemplo, cáncer de estómago, cáncer de hígado, cáncer colorrectal, cáncer de páncreas y cáncer de cuello uterino) para los que existe una asociación causal con el microbioma[88]. Además, la inmunidad innata desempeña un papel fundamental en la evolución de aquellos cánceres con un origen vinculado con la infección, como puede ser la leucemia. Es por ello por lo que, la desregulación de la inmunidad innata es característica de al menos un tercio de todos los cánceres asociados a una inflamación crónica que favorece la aparición de tumores. Se ha descrito que receptores inmunitarios innatos como los TLRs surgen como reguladores críticos de la respuesta inmunitaria a la infección microbiana, desempeñando un papel fundamental en la modulación de una amplia gama de respuestas celulares inhibidoras y promotoras de tumores[88].

Por este motivo, podría ser interesante centrarse en la vía de la inmunidad innata mediada por la señalización TLR y *Myd88* como método para restaurar la respuesta inmunitaria normal en ratones *Pax5*$^{+/-}$.

## 2.1. La alteración del microbioma intestinal promueve el desarrollo de la leucemia en ratones *Pax5*$^{+/-}$, pero no en otras cepas genéticamente predispuestas.

Actualmente, muchos datos apoyan la hipótesis de que el evento que origina el desarrollo de la leucemia infantil es la exposición a la infección. Sin embargo, la naturaleza y el orden de los acontecimientos por los que la infección desencadena la leucemia siguen siendo objeto de debate y hay dos hipótesis que tratan de explicar dicho fenómeno.

Una de estas hipótesis es la teoría de la "infección retardada" la cual postula que la falta de una modulación inmunitaria adecuada en los periodos neonatal e infantil conlleva el riesgo de desarrollar leucemia debido a una exposición retardada a las infecciones comunes en la infancia. De este modo, un sistema inmunitario no expuesto a las infecciones en las primeras etapas del desarrollo puede funcionar de forma anómala cuando se expone posteriormente a dichas infecciones comunes. Además, una señalización inmunitaria anómala desencadenada por las infecciones puede dar lugar a una serie de acontecimientos genéticos que desembocan en el desarrollo completo de la leucemia[30,104].

Alternativamente y simultáneamente, se postuló la hipótesis de "la mezcla de poblaciones" para explicar la aparición de pequeños grupos de pacientes en un lugar y momento localizados[27]. Esta hipótesis expone que la LLA-B infantil podría surgir como resultado infrecuente de la exposición a cualquier infección común, pero que sólo se manifestaría en épocas de mezcla de poblaciones, cuando un número relativamente elevado de niños susceptibles a esta infección se mezclaran con un gran número de individuos infectados, dando lugar a una epidemia generalizada[27].

En cualquiera de los dos casos, una deficiencia transitoria del sistema inmunitario podría actuar, junto con una disbiosis del microbioma intestinal, en las primeras etapas de la vida. Recientemente, se descubrió que tanto los ratones *Pax5*$^{+/-}$ como los *Sca1-ETV6-RUNX1* albergan microbiomas intestinales específicos de cada genotipo directamente relacionados con la alteración genética que portan[100]. Este trabajo estableció un paradigma experimental para evaluar e identificar los factores que modulan la ventaja selectiva de las células preleucémicas. La mayoría de otros tipos de leucemia son independientes de la presencia de infecciones para su progresión. De esta manera, nos preguntamos si la relación "alteración genética-disbiosis específica del microbioma intestinal" es una característica inherente a aquellas condiciones genéticas preleucémicas asociadas a leucemias impulsadas por infecciones o, por el contrario, también está presente en las alteraciones genéticas vinculadas a leucemias independientes de las infecciones.

Para resolver dicha pregunta, se utilizaron dos modelos de ratón que desarrollan leucemia de manera independiente a la exposición a infecciones comunes, el modelo *Sca1-Lmo2*, el cual desarrolla LLA-T[105], y el modelo *Sca1-BCR-ABLp190*, que desarrolla LLA-B[106]. Inicialmente, se analizó el impacto de estos dos genotipos sobre el microbioma utilizando muestras de heces de ratones *Sca1-Lmo2* y *Sca1-BCR-ABLp190*. Se vio que el efecto de cada alteración genética conformaba una composición característica del microbioma intestinal que era distinta entre estas dos cepas de ratones y que también difería del microbioma de los ratones WT (**Figura 18A-B y ANEXO 1**). Del mismo modo, también se analizaron muestras de heces de ratones *Pax5*$^{+/-}$ para corroborar los datos previos obtenidos en los que se indicaba que estos ratones mostraban un microbioma intestinal específico. De esta manera, cada predisposición genética dio forma a una microbiota intestinal única, independientemente de si están asociadas con el desarrollo de la leucemia dependiente o independiente de la infección.

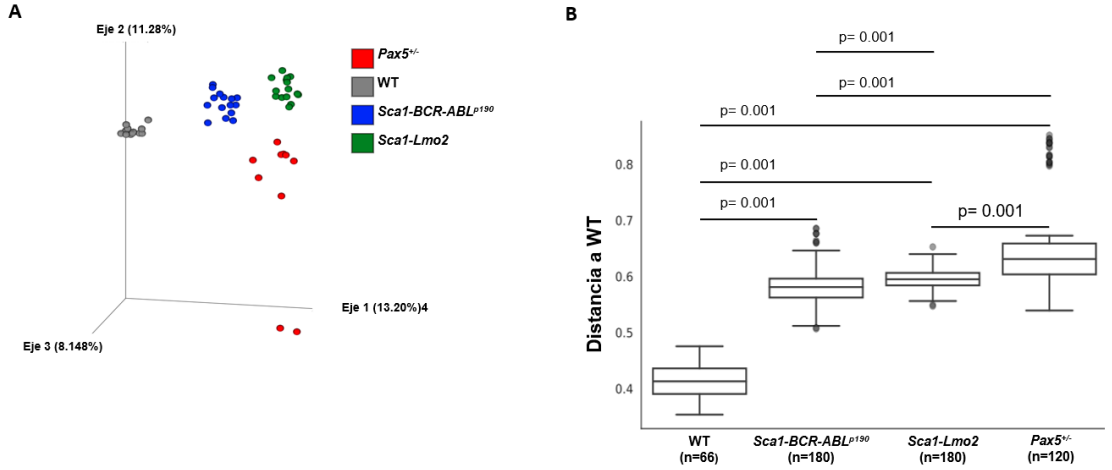



***Figura 18. La susceptibilidad genética a la LLA conforma una microbiota intestinal específica en ratones predispuestos. A)*** *Se calcularon las distancias jaccard por pares (diversidad beta, que refleja cómo de distintas son las muestras comparando unas con otras) para todas las muestras. La métrica de la distancia se ordenó en 3D mediante el análisis de componentes principales (PCA) y se visualizó mediante Emperor. Los ejes indican el porcentaje de varianza explicada.* ***B)*** *Se aplicaron pruebas de Permanova por pares con 999 permutaciones para comprobar las diferencias en la diversidad beta agrupada por genotipo de ratón (WT n=12, Sca1-BCR-ABLp190 n=15, Sca1-Lmo2 n=15 y Pax5$^{+/-}$ n=10). Los recuadros representan las distancias jaccard entre muestras del mismo genotipo. "n=" indica las distancias medidas, "p" es el p-valor de la prueba de Permanova.*

Datos anteriores demostraron que la alteración del microbioma intestinal mediante la depleción bacteriana debida al tratamiento con antibióticos desencadenaba la leucemia en ratones *Pax5*$^{+/-}$ genéticamente predispuestos, incluso en ausencia de infecciones. Basándonos en esta premisa, se decidió explorar si alterar el microbioma intestinal en condiciones SPF podría conducir a la inducción de la leucemia en ratones *Sca1-ETV6-RUNX1*, *Sca1-Lmo2* y *Sca1-BCR-ABLp190*. Estos ratones fueron tratados con un cóctel de antibióticos (ampicilina, vancomicina, ciprofloxacina, imipenem y metronidazol) disuelto en el agua de los biberones durante 8 semanas, iniciándose en torno a las 8 semanas de edad. En los tres modelos, la incidencia de la leucemia y el tiempo transcurrido hasta la aparición de la enfermedad no se vieron alterados por el tratamiento con antibióticos en comparación con los ratones que no siguieron dicho tratamiento (**Figura 19A-B**). Estos resultados ponen de manifiesto que la modulación del microbioma intestinal mediante la eliminación bacteriana afecta exclusivamente al desarrollo de la LLA-B en los ratones *Pax5*$^{+/-}$.

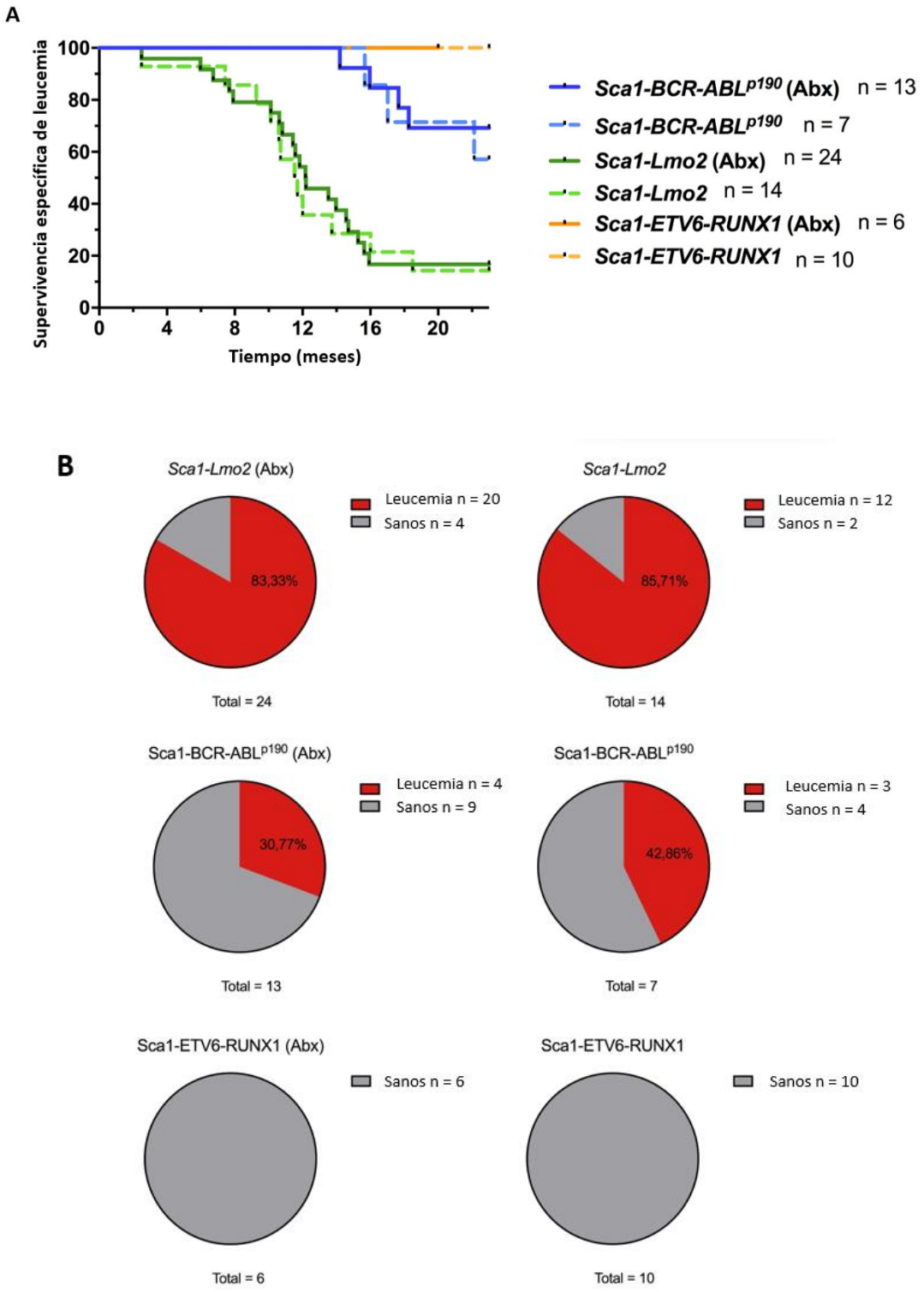



***Figura 19. Curva de supervivencia específica e incidencia de leucemia de ratones Sca1-BCR-ABLp190, Sca1-Lmo2 y Sca1-ETV6-RUNX1 con y sin tratamiento con antibióticos. A)*** *Curva de supervivencia específica de LLA-B en ratones Sca1-Lmo2 tratados (n=24) o no tratados (n=14) con antibióticos (p-value=0.7935), Sca1-BCR-ABLp190 tratados (n=13) o no tratados (n=7) con antibióticos (p-value=0.6514) y Sca1-ETV6-RUNX1 tratados (n=6) o no tratados (n=10) con antibióticos (p-value>0.9999). Ninguno de ellos fue expuesto a infecciones comunes.* ***B)*** *Incidencia de leucemia en ratones Sca1-BCR-ABLp190, Sca1-Lmo2 y Sca1-ETV6-RUNX1 con y sin privación de la microbiota intestinal. Todos los ratones fueron alojados en condiciones SPF y algunos de los ratones de los tres genotipos fueron tratados con un cóctel de antibióticos durante 8 semanas para inducir la alteración del microbioma intestinal. No se encontraron diferencias significativas mediante la prueba exacta de Fisher. En total, ninguno de los ratones Sca1-ETV6-RUNX1 privados del*

*microbioma (n=6) desarrolló LLA-B; el 30,77% (4/13) de los Sca1-BCR-ABLp190 tratados con un antibiótico desarrollaron LLA-B y el 83.33% (20/24) de los Sca1-Lmo2 privados del microbioma desarrollaron LLA-T. Los ratones Sca1-ETV6-RUNX1 necesitan exposición a la infección para desarrollar B-ALL, por esta razón, ninguno de los ratones desarrolló B-ALL en estas condiciones de alojamiento. Abx: tratados con un cóctel de antibióticos durante 8 semanas.*

## 2.2. El entrenamiento inmunitario mediante la exposición temprana a las infecciones no previene el desarrollo de la leucemia en ratones *Pax5*$^{+/-}$.

De acuerdo con la hipótesis de la exposición retardada, los estudios epidemiológicos sugieren que las exposiciones que conducen a la estimulación inmunológica (lactancia materna, asistencia a guarderías, contacto con mascotas, alergias, etc.) durante la infancia están inversamente correlacionadas con el desarrollo de los subtipos más comunes de LLA-B (*ETV6-RUNX1+* y LLA-B de alta hiperdiploidía)[29,107,108,109,110,111]. Estos hallazgos plantean la posibilidad de que la falta de exposición a infecciones durante los primeros años de vida contribuya a la aparición de la LLA-B. Hay algunas observaciones que sugieren que este podría no ser el caso del desarrollo de la leucemia en los ratones *Pax5*$^{+/-}$, como es el hecho de que la LLA-B también aparece cuando los ratones de más edad (> 5 meses) se exponen a infecciones naturales[25] o que la eliminación de las células B preleucémicas mediante la inhibición transitoria de la vía JAK/STAT en el momento de la exposición a la infección previene el desarrollo de la LLA-B[112]. Apoyando dichas observaciones, recientes estudios epidemiológicos en humanos demuestran que no hay evidencias de que la exposición temprana a infecciones comunes, debido a la asistencia a guarderías, proteja frente al desarrollo de leucemia en niños[113].

Con el fin de determinar si el momento de exposición a la infección es o no relevante para el desarrollo de la LLA-B, se expusieron ratones *Pax5*$^{+/-}$ y ratones WT a las mismas condiciones infecciosas desde el embarazo y el nacimiento, y después durante el resto de su vida (este grupo se denomina “grupo de exposición temprana”). De igual modo, otro grupo de ratones *Pax5*$^{+/-}$ y WT nacieron y se mantuvieron hasta los 2 meses de edad en condiciones SPF, momento en el cual fueron trasladados a una instalación convencional de exposición a las infecciones (este grupo se denomina “grupo de

exposición tardía") (**Figura 20**). El estado microbiológico en la instalación convencional se definió y controló durante 2 años (**ANEXO 2**).

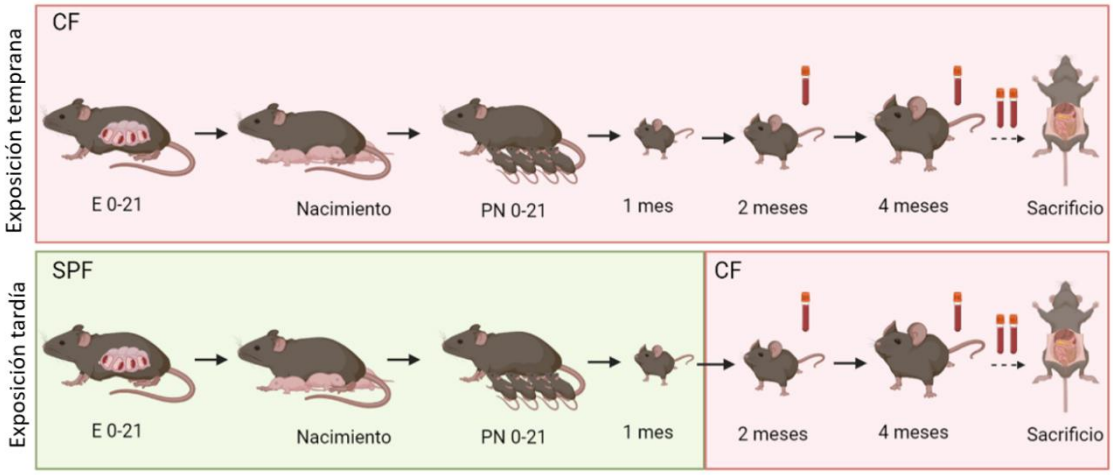



***Figura 20. Diseño experimental del estudio del entrenamiento inmune en ratones Pax5$^{+/-}$.*** *Las hembras de ratón del "grupo de exposición temprana" quedaron preñadas en una instalación convencional (CF) donde están expuestas a infecciones comunes durante toda su vida, y las crías nacidas se mantuvieron en estas condiciones durante toda su vida. Las hembras del grupo de "exposición tardía" quedaron preñadas en condiciones SPF, y las crías se mantuvieron en SPF hasta el mes de edad, cuando se trasladaron a las instalaciones convencionales. Los ratones fueron controlados periódicamente. El punto final se estableció a los 2 años de edad (22-24 meses) o cuando los ratones mostraron signos de enfermedad (signos clínicos o presencia de blastos en sangre periférica). Los animales fueron sacrificados y se analizaron los tejidos hematopoyéticos y no hematopoyéticos. E: periodo embrionario; PN: periodo postnatal; SPF: libre de patógenos específicos.*

Bajo estas condiciones, en el grupo de exposición temprana se observó un desarrollo de la LLA-B en el 25% (7 de 28) de los animales *Pax5$^{+/-}$*, pero no en los ratones WT (**Figura 21A**), datos que se asemejan mucho a la penetrancia de LLA-B tanto en el grupo de exposición tardía[25] como en las familias con la mutación heterocigota *PAX5* c.547G>A[114,115,116]. La aparición de la leucemia en el grupo de exposición temprana a las infecciones se produjo entre los 9 y los 17 meses de edad, de forma similar al grupo de exposición tardía[25] y con el mismo fenotipo leucémico (**Figura 21B**). Por lo tanto, estos resultados indican que el momento de la infección no afecta al desarrollo de la leucemia en el modelo *Pax5$^{+/-}$* y que la presencia de un estrés inmunitario (el alojamiento en una instalación convencional) es el acontecimiento desencadenante.

No obstante, todavía era una incógnita si los mecanismos que provocaban la susceptibilidad a la LLA-B en los ratones *Pax5$^{+/-}$* sometidos a una exposición temprana a las infecciones eran similares a los que provocaban la LLA-B en el grupo *Pax5$^{+/-}$* de exposición tardía. Se estudiaron las células B presentes en los ratones *Pax5$^{+/-}$* expuestos

de manera temprana a las infecciones y se observó una cantidad significativamente reducida de células B totales en la sangre periférica en comparación con las que presentaban los ratones WT en las mismas condiciones (**Figura 22**). Esta disminución de las células B en la SP fue similar a la observaba en el grupo de ratones *Pax5*$^{+/-}$ de exposición tardía (**Figura 22**). De esta manera, la reducción del porcentaje de células B en SP es debida a la alteración genética, ratones *Pax5*$^{+/-}$ con predisposición a la LLA-B en este caso, y no al momento en el que estos animales han sido expuestos al estímulo infeccioso.

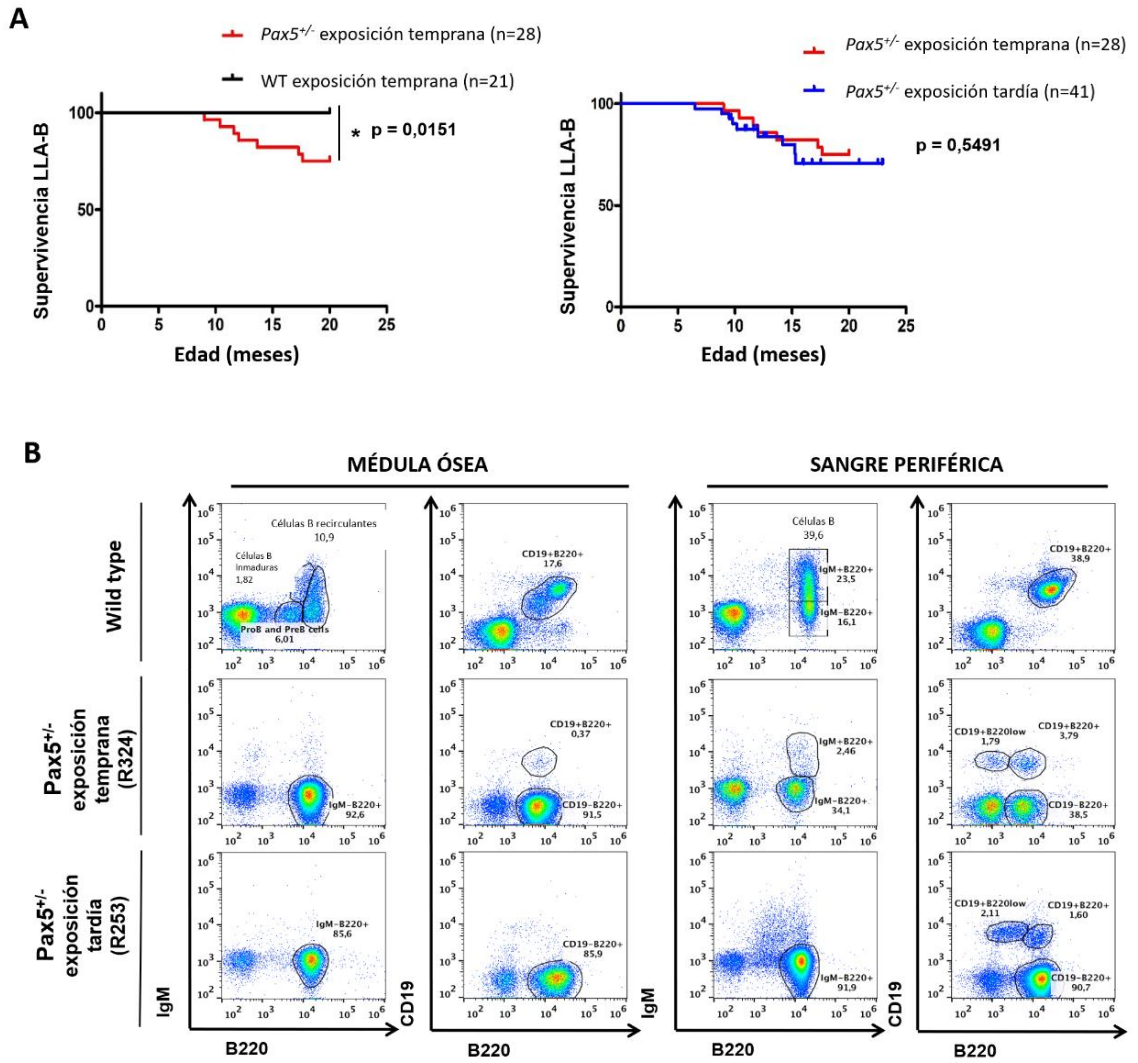



***Figura 21. La exposición temprana a la infección promueve el desarrollo de la leucemia linfoblástica aguda de células B precursoras en ratones Pax5$^{+/-}$. A)*** *Curva de supervivencia específica de la LLA-B en los animales Pax5$^{+/-}$ nacidos y mantenidos en CF (rojo, n=28) en comparación con los ratones de control WT (negro, n=21) mostrando una esperanza de vida significativamente más corta (p-valor=0,0151). La curva de supervivencia específica de la LLA-B del grupo de exposición temprana Pax5$^{+/-}$ (rojo, n=28) en comparación con el grupo de exposición tardía Pax5$^{+/-}$ (azul, n=41) muestra una esperanza de vida similar (p-valor=0,5491) (panel derecho).* ***B)*** *Análisis de citometría de flujo de poblaciones hematopoyéticas en ratones enfermos Pax5$^{+/-}$ de exposición temprana. Los gráficos que representan las poblaciones celulares de la médula ósea y la sangre periférica muestran una acumulación de blastos B en los ratones*

*Pax5$^{+/-}$ expuestos de manera temprana (n=7, edad: 9-17 meses) en comparación con los ratones WT de control de la misma edad (n=4, edad: 8-16 meses). Los ratones Pax5$^{+/-}$ de exposición temprana muestran un fenotipo inmunitario similar al de los ratones Pax5$^{+/-}$ expuestos tardíamente (n=9, edad: 6-16 meses).*

A continuación, se llevó a cabo la secuenciación del exoma completo (WES, de sus siglas en inglés *whole exome sequencing*) de 4 tumores de ratones *Pax5$^{+/-}$* de exposición temprana a la infección con el fin de identificar los eventos secundarios adquiridos somáticamente y que conducen al desarrollo de la leucemia. De forma similar a las leucemias que presentan los ratones *Pax5$^{+/-}$* de exposición tardía, los tumores de los ratones expuestos tempranamente mostraron mutaciones recurrentes que afectaban al alelo *Pax5* WT restante, como la mutación *Pax5$^{P80R}$* que estaba presente en 2 de las 4 leucemias analizadas (**Tabla 1 y ANEXO 3**). Además, la mutación *Jak1$^{F837V}$* también apareció en una de las muestras leucémicas, resaltando de nuevo la importancia de la vía de señalización JAK/STAT en las leucemias *Pax5$^{+/-}$*, como se había demostrado anteriormente con las mutaciones *Jak1* y *Jak3* en las leucemias de los ratones *Pax5$^{+/-}$* de exposición tardía a la infección[25,112]. Curiosamente, 2 de los 4 tumores analizados presentaban mutaciones somáticas que afectaban a la vía RAS-RAF-MEK-ERK (*Nras$^{G13D}$* y *Kras$^{Q61H}$*) (**Tabla 1 y ANEXO 3**). Las mutaciones somáticas que desregulan esta vía de señalización constituyen una de las aberraciones genéticas más comunes en la LLA-B infantil humana[117].

En resumen, se han obtenido pruebas *in vivo* que demuestran que los ratones *Pax5$^{+/-}$* expuestos a infecciones en etapas tempranas de la vida desarrollan leucemia con una incidencia, latencia y características moleculares similares a las de los ratones *Pax5$^{+/-}$* expuestos en etapas posteriores de la vida. Además, estos resultados muestran que el entrenamiento inmunitario mediante la exposición temprana a las infecciones no tiene un efecto protector contra la conversión de células preleucémicas a leucémicas en los ratones heterocigotos para una mutación de *Pax5* en la línea germinal. Estos hallazgos sugieren un defecto celular autónomo dentro de las células preleucémicas *Pax5$^{+/-}$* y refuerzan aún más la visión de la variación etiológica en el desarrollo de la leucemia entre los diferentes subtipos moleculares de LLA-B.

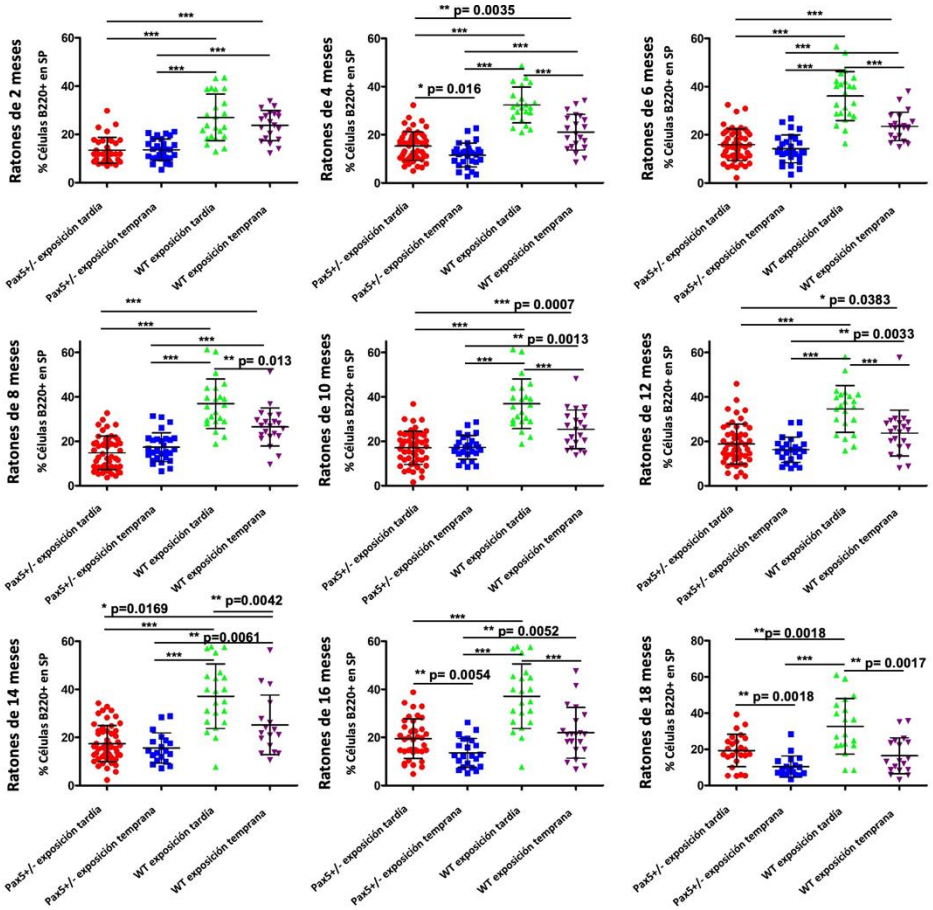



***Figura 22. Disminución de células B presentes en la SP de ratones Pax5$^{+/-}$.*** *Porcentaje de células B en sangre periférica (B220$^{+}$IgM$^{+/-}$) en diferentes puntos temporales en ratones Pax5$^{+/-}$ de exposición temprana (n=19-28) comparados con ratones WT de exposición temprana (n=19-21), ratones Pax5$^{+/-}$ de exposición tardía (n=25-34) y ratones WT de exposición tardía (n=18-23) analizados por citometría de flujo. Las barras de error representan la desviación estándar. Los p-valores corresponden a los obtenidos por un test t no pareado.*

***Tabla 1. Mutaciones secundarias en ratones Pax5$^{+/-}$ portadores de tumores expuestos tempranamente a la infección.*** *Se realizó la secuenciación del exoma completo de 4 tumores de exposición temprana de ratones Pax5$^{+/-}$ y de la línea germinal correspondiente en una plataforma HiSeq 2500 (Illumina). El ADN de los tumores de exposición temprana Pax5$^{+/-}$ procedía de la médula ósea leucémica, mientras que se utilizó como material de referencia de la línea germinal el ADN de la cola del ratón. Del mismo modo que las leucemias Pax5$^{+/-}$ del grupo de exposición tardía, los tumores Pax5$^{+/-}$ de exposición temprana mostraron mutaciones recurrentes que afectaban al alelo Pax5 WT restante, con la mutación Pax5$^{P80R}$ presente en 2 de las 4 leucemias analizadas. Todas las variantes encontradas por WES fueron confirmadas por secuenciación Sanger.*

| | Número de proyecto de secuenciación de ADN | ID | Edad de enfermedad del ratón | Fenotipo *CD19* | Mutación *JAK3* en ratón | Mutación *JAK3* homóloga en humano | Mutación *Pax5* en el alelo WT | Mutación *Pax5* homóloga en humano | Otras mutaciones |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 123BM | R324 | 10,4 meses | CD19- | - | - | - | - | - |
| 2 | 141BM | R330 | 12,03 meses | CD19+ | - | - | P80R | P80R | *Kras* Q61H |
| 3 | 157BM | T830 | 9 meses | CD19 bajo | - | - | P80R | P80R | *Jak1* F837V |
| 4 | 158T | R017 | 17,23 meses | CD19- | - | - | - | - | *Nras* G13D |

## 2.3. El estrés inmunológico disminuye la expresión de *Myd88* en células preleucémicas *Pax5*$^{+/-}$.

Como se ha mencionado anteriormente, la inmunidad innata desempeña un papel fundamental en la evolución de aquellos cánceres que tienen una asociación causal demostrada con la infección[88]. De hecho, se ha observado que, en los cánceres desencadenados por infecciones, los TLRs desempeñan un papel fundamental en la modulación de las respuestas celulares inhibidoras y promotoras de tumores, tanto a nivel de las células inmunitarias del microambiente tumoral como actuando directamente sobre las propias células cancerosas[88].

Para dilucidar si la señalización de los TLRs es necesaria en la evolución clonal de las células B premalignas en aquellas leucemias asociadas a la infección, se realizó un análisis mediante PCR (reacción en cadena de la polimerasa, del inglés *polymerase chain reaction*) de los genes implicados en la vía de señalización de los TLRs. Para ello se utilizaron precursores de células B aislados de la médula ósea (MO) de ratones *Pax5*$^{+/-}$ y WT mantenidos tanto en condiciones SPF como expuestos a las infecciones comunes durante 3 y 6 meses. Como se había demostrado anteriormente, los ratones *Pax5*$^{+/-}$ sólo desarrollaron LLA-B bajo la exposición a la infección[25]. El análisis de la vía de señalización TLR en las células B precursoras preleucémicas de estos ratones *Pax5*$^{+/-}$ que habían desarrollado la leucemia, reveló una disminución de la expresión de los adaptadores *Myd88* y *Tirap*, así como del factor de transcripción *Nfkb1a* (**Figura 23A**).

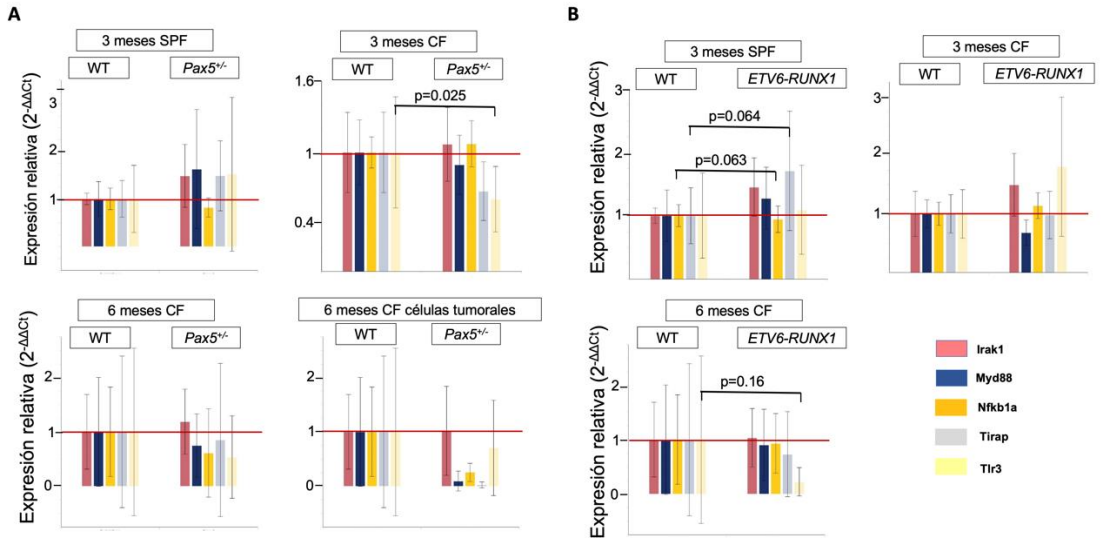


***Figura 23. Perfil del análisis mediante PCR de la vía de señalización de los TLRs de ratón. A)*** *Se evaluaron mediante un array de la vía de señalización de los TLRs las células proB (B220$^{bajo}$ IgM$^{-}$) aisladas de la MO de ratones WT (n= 4), ratones sanos Pax5$^{+/-}$ (n=4) y ratones Pax5$^{+/-}$ leucémicos (n= 2) expuestos a diferentes ambientes (SPF, instalación libre de patógenos y CF, instalación convencional).* ***B)*** *Mismo análisis que el anterior, pero en esta ocasión con células proB (B220$^{bajo}$ IgM$^{-}$) aisladas de la MO de ratones WT (n=4) y ratones Sca1-ETV6-RUNX1 sanos (n=4) expuestos a diferentes contextos ambientales (instalación SPF y CF). Los datos se representan como expresión relativa (2-ΔΔΔCt) utilizan el método comparativo CT. Los p-valores se calculan a partir de una prueba t de Student de los valores 2$^{(-\ Delta\ CT)}$ replicados para cada gen.*

Además, para confirmar los resultados observados con el array de los TLRs, también se llevó a cabo una PCR cuantitativa en tiempo real, específica para *Myd88*, con la que se observó que un porcentaje próximo al 30% (1 de 3) de los ratones *Pax5*$^{+/-}$ presentaban bajos niveles de expresión de *Myd88* en las células B preleucémicas. Este dato resulta interesante ya que este porcentaje se asemeja al 22% de incidencia de LLA-B que previamente se había visto en los ratones *Pax5*$^{+/-}$[25], sugiriendo que las células que ven reducidos sus niveles de *Myd88* son aquellas que posteriormente se transforman en leucémicas (**Figura 24**). Curiosamente, al realizar este mismo análisis de la expresión de *Myd88* en las células B preleucémicas de ratones *Pax5*$^{+/-}$, pero en esta ocasión, tratados con antibióticos, se observó una reducción significativa y homogénea de la expresión de *Myd88,* siendo estos niveles de *Myd88* similares a los que presentan las células B leucémicas de ratones *Pax5*$^{+/-}$ tratados también con antibióticos (**Figura 24**). De esta manera, se confirma el hecho de que tanto las células preleucémicas como las propias células ya leucémicas presentan bajos niveles de expresión de *Myd88*.

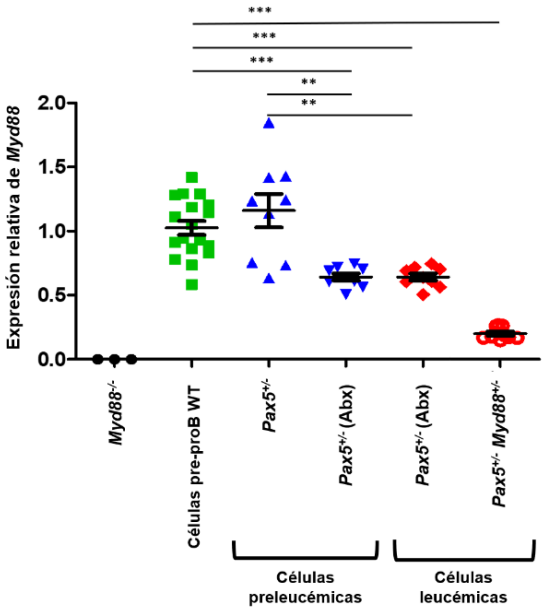


***Figura 24. El estrés inmunológico mediado por la exposición a infecciones comunes o por el tratamiento con antibióticos disminuye la expresión de Myd88 en células preleucémicas Pax5$^{+/-}$.*** *Cuantificación por PCR en tiempo real (q-PCR) de los niveles de expresión de Myd88 en células proB Pax5$^{+/-}$ preleucémicas y leucémicas. Muestras preleucémicas: células proB (B220$^{bajo}$ IgM$^{-}$) aisladas de MO de ratones WT de 3 meses de edad (n=5), Pax5$^{+/-}$ (n=3) y Pax5$^{+/-}$ tratados con antibióticos (Abx) durante 8 semanas (n=3). Muestras leucémicas: células leucémicas de la MO de ratones Pax5$^{+/-}$ tratados con Abx durante 8 semanas (n=3) y ratones Pax5$^{+/-}$;Myd88$^{+/-}$ (n=3). Todos los ratones analizados por q-PCR fueron expuestos a infecciones. La médula ósea total de los ratones Myd88$^{-/-}$ se utilizó como control negativo. (*** p-valor<0,0001; ** p-valor=0,0014).*

Como se ha mencionado previamente, los genotipos *Pax5$^{+/-}$* y *ETV6-RUNX1* son modelos característicos de leucemias asociadas a la infección. Por ello, también se estudió la expresión de *Myd88* en ratones *Sca1-ETV6-RUNX1*. El análisis reveló que, igual que en el caso anterior con los ratones *Pax5$^{+/-}$*, también se producía una disminución de la expresión de *Myd88* en las células B preleucémicas de la médula ósea (**Figura 23B**). Cabe destacar que también se observaron niveles reducidos de la expresión de *Myd88* en células leucémicas humanas *ETV6-RUNX1+* en comparación con las células de médula ósea de donantes sanos (**Figura 25**).

Estos resultados reflejan que la inducción de un estado leucémico se asocia con una disminución de la expresión de *Myd88* en las células B progenitoras tanto humanas como de ratón.

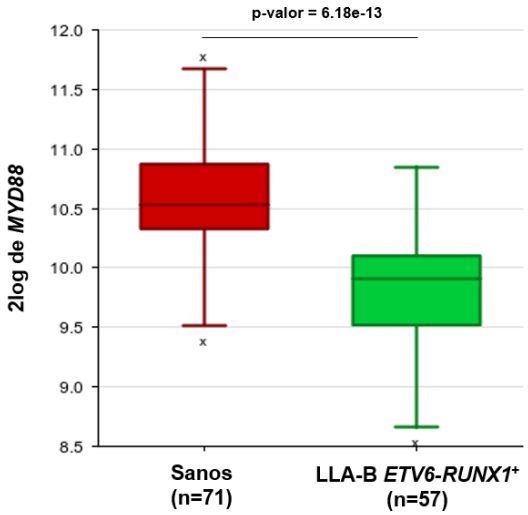
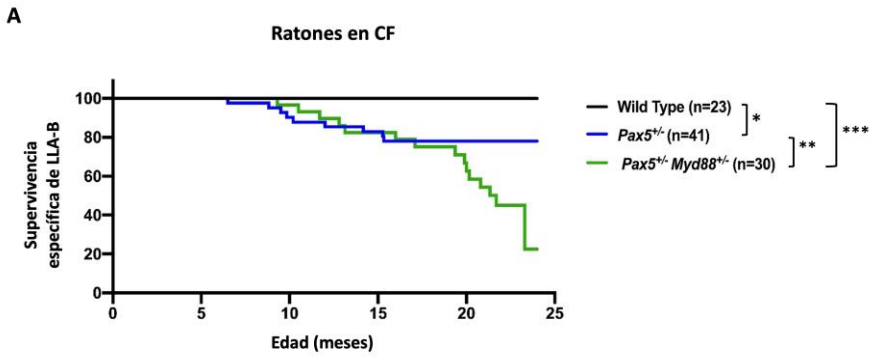



***Figura 25. La expresión de MYD88 está disminuida en las leucemias humanas ETV6-RUNX1+.*** *Utilizando la plataforma de análisis y visualización genómica R2 (http://r2.amc.nl) se analizó la expresión génica de MYD88 en muestras leucémicas humanas de MO ETV6-RUNX1+ (n=57) y en muestras de MO sana (n=71) del conjunto de datos MILE (Microarray Innovations in LEukemia) disponible en dicha plataforma[118]. El p-valor = 6.18e-13. "x" representa un valor atípico.*

## 2.4. La reducción en la expresión de *Myd88* aumenta la incidencia de LLA-B en ratones *Pax5*$^{+/-}$ expuestos a infecciones comunes.

Dada la reducción en la expresión de *Myd88* observada tanto en las células preleucémicas como en las leucémicas de los ratones *ETV6-RUNX1* y *Pax5*$^{+/-}$, a continuación, se quiso comprobar directamente la necesidad de *Myd88* en el desarrollo de la LLA-B. En primer lugar, se disminuyó la expresión de *Myd88* generando ratones *Pax5*$^{+/-}$*;Myd88*$^{+/-}$ para comprobar si dichos ratones eran más propensos a desarrollar LLA-B asociada a la infección que los ratones *Pax5*$^{+/-}$. Los ratones *Pax5*$^{+/-}$*;Myd88*$^{+/-}$ fueron expuestos a infecciones comunes y se monitorizaron para detectar cualquier signo de enfermedad.

Los resultados mostraron que el desarrollo de la LLA-B era significativamente mayor en los ratones *Pax5*$^{+/-}$*;Myd88*$^{+/-}$, en los que el 64% (16 de 25) terminaron desarrollando leucemia (**Figura 26A**), muy similar a la incidencia presentada por los ratones *Pax5*$^{+/-}$ tratados con antibióticos donde la incidencia era del 63%[32]. La aparición de la leucemia en estos ratones *Pax5*$^{+/-}$*;Myd88*$^{+/-}$ se produjo entre los 7 y los 20 meses de edad (**Figura 26A**), de forma también análoga al grupo de ratones *Pax5*$^{+/-}$. Además, se observó que un 62,5% (5 de 8) de los ratones *Pax5*$^{+/-}$*;Myd88*$^{+/-}$ desarrollaron LLA-B en ausencia de exposición a la infección, condición en la que los ratones *Pax5*$^{+/-}$ nunca presentan la enfermedad[25] (**Figura 26B)**. Este hecho permitió confirmar el papel crítico que tiene la disminución de la expresión de *Myd88* en el desarrollo de la LLA-B en los ratones *Pax5*$^{+/-}$. En todos estos casos, la caracterización de las leucemias de los ratones *Pax5*$^{+/-}$*;Myd88*$^{+/-}$ mostraron que eran fenotípicamente similares a las que aparecen en el grupo de los *Pax5*$^{+/-}$ (**Figura 27A-B**).



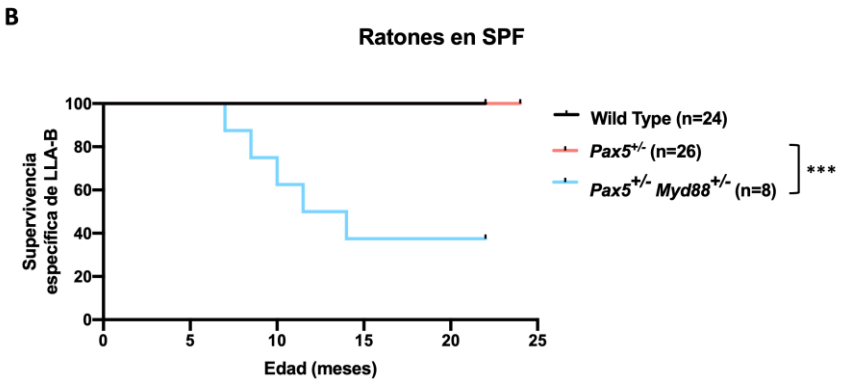



***Figura 26. Supervivencia específica de la LLA-B en ratones Pax5$^{+/-}$ y Pax5$^{+/-}$;Myd88$^{+/-}$. A)*** *Se representa la supervivencia específica de leucemia de los siguientes genotipos expuestos a infecciones: Pax5$^{+/-}$ (línea azul, n=41); Pax5$^{+/-}$;Myd88$^{+/-}$ (línea verde, n=30); y WT (línea negra, n=23). p-valor = 0,0178 en la prueba de Mantel-Cox al comparar ratones Pax5$^{+/-}$ con ratones WT; p-valor <0,0001 al comparar ratones Pax5$^{+/-}$;Myd88$^{+/-}$ con ratones WT; y p-valor = 0,0092 cuando se comparan ratones Pax5$^{+/-}$ frente a ratones Pax5$^{+/-}$;Myd88$^{+/-}$.* ***B)*** *Se representa la supervivencia específica de leucemia de los siguientes genotipos en ausencia de exposición a infecciones: Pax5$^{+/-}$; Myd88$^{+/-}$ (línea azul, n=8); Pax5$^{+/-}$ (línea roja, n=26); y WT (línea negra, n=24). p-valor <0,0001 en la prueba de Mantel-Cox al comparar ratones Pax5$^{+/-}$;Myd88$^{+/-}$ y Pax5$^{+/-}$.*

Para explorar la relevancia que tienen estos hallazgos previos de modelos murinos en la leucemia humana, se comparó el perfil de expresión de las leucemias de ratones *Pax5$^{+/-}$;Myd88$^{+/-}$* con las células tumorales de pacientes con LLA-B[119,120,121]. Mediante el análisis de la expresión génica con microarrays de ARN se identificaron genes expresados diferencialmente (n=2463) en la MO de ratones *Pax5$^{+/-}$;Myd88$^{+/-}$* leucémicos en comparación con las células B precursoras presentes en la MO de ratones WT (**Figura 28A**). Cuando se compararon con firmas previamente definidas obtenidas a partir del análisis de leucemias humanas, estos genes expresados de forma diferencial en los ratones estaban enriquecidos en perfiles de expresión específicos de las LLA-B humanas[119,120,122], lo que pone de manifiesto las similitudes entre las leucemias humanas y murinas (**Figura 28B**). No obstante y como era de esperar, se observó que las leucemias de los ratones *Pax5$^{+/-}$;Myd88$^{+/-}$* mostraban un patrón de expresión génica específico (**Figura 29A-B**).

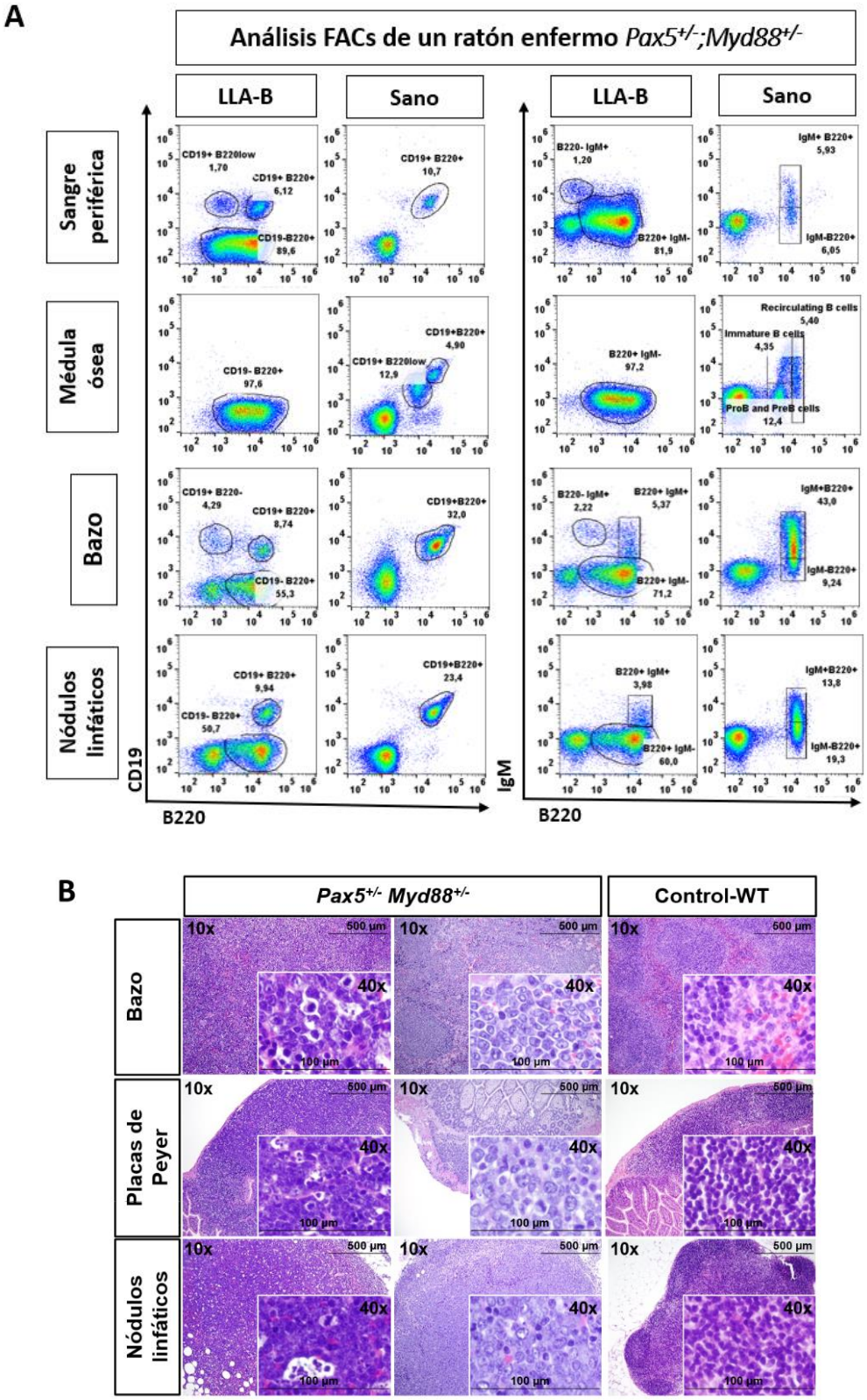
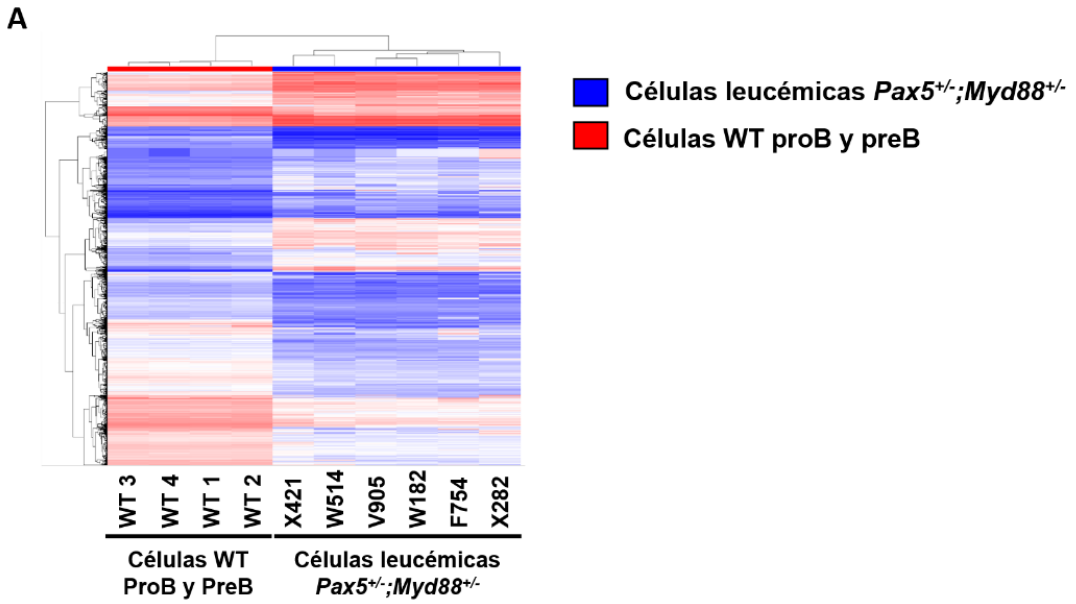



***Figura 27. Ejemplo de LLA-B en un ratón Pax5$^{+/-}$;Myd88$^{+/-}$. A)*** *Análisis de citometría de flujo que muestra la acumulación de células blásticas (B220$^{+}$ IgM$^{-}$ CD19$^{-}$) en sangre periférica, médula ósea, bazo y nódulos linfáticos en un ratón leucémico Pax5$^{+/-}$;Myd88$^{+/-}$ (W182). Se compara con un ratón Pax5$^{+/-}$;Myd88$^{+/-}$ sano de la misma edad (M706), ambos expuestos a infecciones comunes.* ***B)*** *Tinción con hematoxilina y eosina de ratones Pax5$^{+/-}$;Myd88$^{+/-}$ portadores de tumores que muestran células blásticas infiltrantes en el bazo, placas de Peyer y ganglios linfáticos, en comparación con un ratón WT de la misma edad. Se aprecia pérdida de la arquitectura normal por infiltración de células leucémicas. En cada caso se indican los aumentos y la barra de escala correspondiente.*

A continuación, se quisieron identificar los eventos secundarios que se adquieren somáticamente y que conducen al desarrollo de la LLA-B en los ratones *Pax5*$^{+/-}$;*Myd88*$^{+/-}$. Para ello, se realizó la secuenciación completa del genoma (WGS, de sus siglas en inglés *whole genome sequencing*) de 8 muestras de ADN tumoral (médula ósea o nódulos linfáticos), así como de sus correspondientes muestras de ADN control obtenidas de la línea germinal de la cola de los ratones *Pax5*$^{+/-}$;*Myd88*$^{+/-}$. El porcentaje de células leucémicas en las muestras secuenciadas osciló entre el 88,1% y el 98,7%.

Los resultados permitieron identificar varias mutaciones recurrentes adquiridas somáticamente y variaciones en el número de copias que afectaban a factores de transcripción de las células B en los ratones *Pax5*$^{+/-}$;*Myd88*$^{+/-}$ enfermos (**Figura 30 y ANEXO 4**). Además, se detectaron mutaciones recurrentes que afectaban a las vías de señalización de JAK/STAT y RAS (**Figura 30**), similares a las identificadas en las leucemias murinas *Pax5*$^{+/-}$[25,100,112] y en las muestras de LLA-B humanas.

Así pues, los impulsores de la LLA-B asociada a la infección son parecidos tanto en las células leucémicas *Pax5*$^{+/-}$;*Myd88*$^{+/-}$ como en las *Pax5*$^{+/-}$. Estas observaciones revelan que la reducción de *Myd88* aumenta la incidencia de la leucemia y lo hace con un perfil genético similar al de las leucemias observadas en los ratones *Pax5*$^{+/-}$.



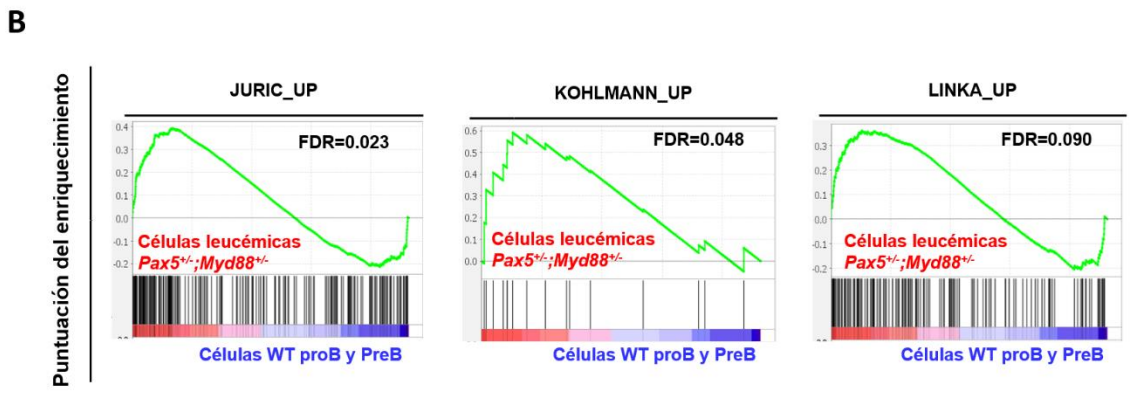



***Figura 28. Análisis de la expresión génica de ratones leucémicos Pax5$^{+/-}$;Myd88$^{+/-}$. A)*** *Heatmap no supervisado que muestra los genes expresados diferencialmente (n=2463) entre células leucémicas de ratones Pax5$^{+/-}$;Myd88$^{+/-}$ (n=6) y células proB y preB de ratones WT (n=4). El análisis de significación de los microarrays se definió por un p-valor < 0,05 y un cambio de expresión: FC< -2 o FC> 2. (FC, del inglés fold change.)* ***B)*** *GSEA que muestra que las células leucémicas Pax5$^{+/-}$;Myd88$^{+/-}$ de ratones enfermos están enriquecidas en conjuntos de genes específicos de LLA-B humanas. El análisis de enriquecimiento se realizó utilizando conjuntos de genes de firmas de expresión génica descritos como infraexpresados o sobreexpresados específicamente en la LLA-B humana[119,120,122]. Los conjuntos de genes mostrados están significativamente enriquecidos con un p-valor < 1%.*

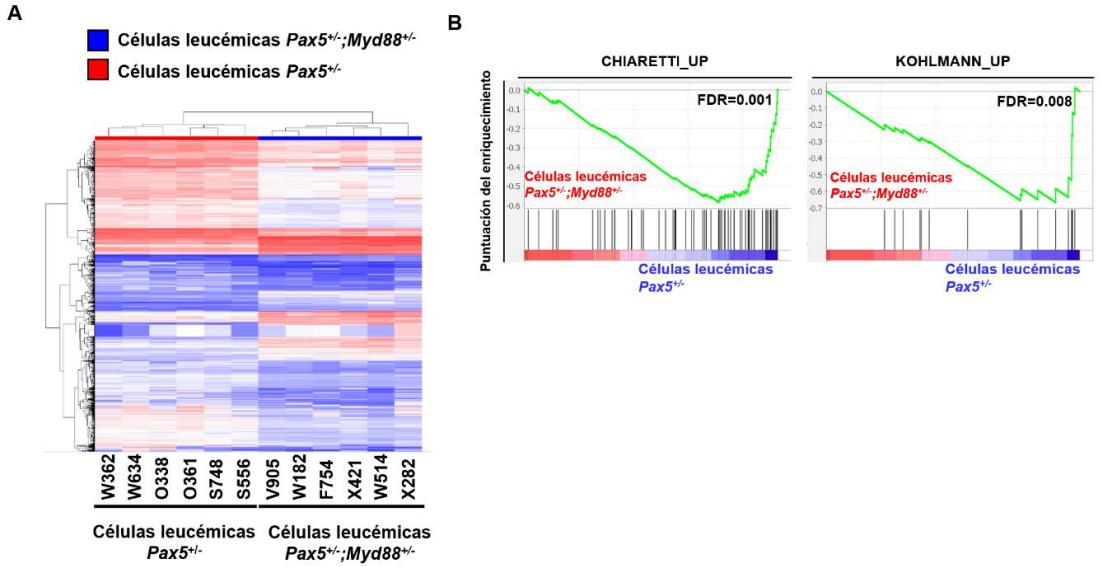



***Figura 29. Análisis de expresión génica comparando muestras de ratones leucémicos Pax5$^{+/-}$;Myd88$^{+/-}$ con ratones leucémicos Pax5$^{+/-}$. A)*** *Heatmap no supervisado que muestra los genes expresados diferencialmente (n=1385) entre las células leucémicas de ratones Pax5$^{+/-}$;Myd88$^{+/-}$ (n=6) y las células leucémicas de ratones Pax5$^{+/-}$ (n=6). El análisis de significación de los microarrays se definió por un p-valor <0,05 y un cambio de expresión: FC< -2 o FC> 2. (FC: fold change).* ***B)*** *GSEA que muestra que las células leucémicas Pax5$^{+/-}$ de ratones enfermos están enriquecidas en conjuntos de genes específicos de LLA-B humanas que no están enriquecidos en ratones leucémicos Pax5$^{+/-}$;Myd88$^{+/-}$.*

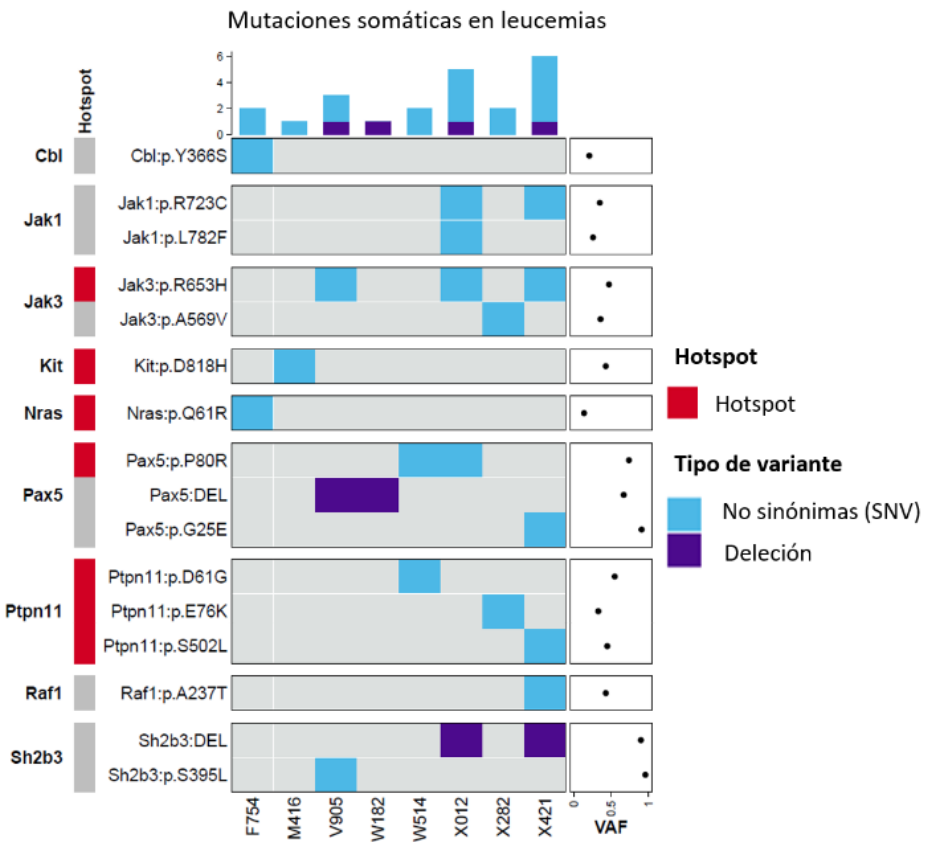



***Figura 30. Secuenciación del genoma completo en LLA-B de ratones Pax5$^{+/-}$;Myd88$^{+/-}$.*** *OncoPrint de las mutaciones somáticas de un solo nucleótido y alteraciones del número de copias encontradas en la secuenciación del genoma completo de las 8 muestras de leucemia analizadas. El ADN tumoral se obtuvo a partir de la médula ósea o de los nódulos linfáticos leucémicos completos, mientras que el ADN de la cola del ratón correspondiente se utilizó como material de la línea germinal de referencia. En la parte inferior se muestra el código identificativo de cada columna. Las alteraciones somáticas se agrupan por genes, ordenados por nombre en la parte izquierda de la figura. En la columna de la izquierda se muestran los genes que se han descrito como mutaciones secundarias en leucemias humanas o de ratón, y coloreados en rojo los hotspot (del inglés, puntos calientes, los sitios del gen más frecuentemente mutados) de los distintos genes. En la columna de la derecha se muestra un mapa de puntos que representa la media de la fracción de alelos con la variante tumoral de cada mutación encontrada (VAF, de las siglas en inglés variant allele fraction). En la leyenda se muestra el color que identifica cada tipo de mutación encontrada; SNV se refiere a mutaciones de nucleótido único (del inglés, single nucleotide variation).*

## 2.5. Las células B precursoras de *Pax5$^{+/-}$;Myd88$^{+/-}$* presentan un perfil transcripcional inflamatorio.

Hasta ahora, en este trabajo se ha descrito cómo las células B preleucémicas *Pax5$^{+/-}$* expuestas a un estrés inmunológico, bien infecciones o bien un tratamiento con antibióticos, sufren una reducción en la expresión de *Myd88*, la cual provoca un aumento en la incidencia de la LLA-B en los ratones *Pax5$^{+/-}$*. A continuación, se quiso explorar la función de esta disminución en la expresión de *Myd88* mediante la

comparación entre los perfiles transcripcionales de las células B precursoras preleucémicas de los ratones *Pax5*$^{+/-}$;*Myd88*$^{+/-}$ y las de los *Pax5*$^{+/-}$ (**Figura 31A-E**).

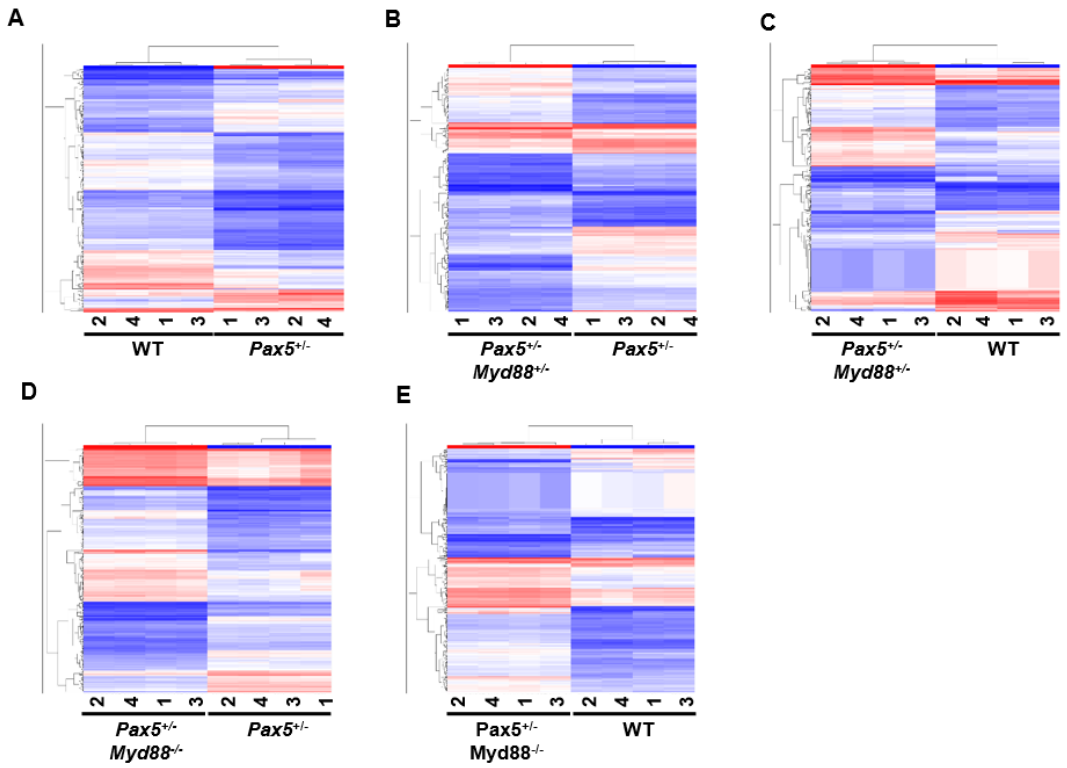



***Figura 31. Análisis de la expresión génica de células proB sanas de ratones WT, Pax5$^{+/-}$, Pax5$^{+/-}$;Myd88$^{+/-}$ y Pax5$^{+/-}$;Myd88$^{-/-}$. A)*** *Heatmap no supervisado que muestra los genes expresados diferencialmente (n=265) entre células proB sanas de ratones Pax5$^{+/-}$ (n=4) y ratones WT (n=4).* ***B)*** *Heatmap no supervisado que muestra los genes expresados diferencialmente (n=237) entre células proB sanas de ratones Pax5$^{+/-}$;Myd88$^{+/-}$ (n=4) y ratones Pax5$^{+/-}$ (n=4).* ***C)*** *Heatmap no supervisado que muestra los genes expresados diferencialmente (n=327) entre células proB sanas de ratones Pax5$^{+/-}$;Myd88$^{+/-}$ (n=4) y ratones WT (n=4).* ***D)*** *Heatmap no supervisado que muestra los genes expresados diferencialmente (n=325) entre células proB sanas de ratones Pax5$^{+/-}$;Myd88$^{-/-}$ (n=4) y ratones Pax5$^{+/-}$ (n=4).* ***E)*** *Heatmap no supervisado que muestra los genes expresados diferencialmente (n=349) entre células proB sanas de ratones Pax5$^{+/-}$;Myd88$^{-/-}$ (n=4) y ratones WT (n=4). El análisis de significación de los microarrays se definió por un P-valor <0,05 y un cambio de expresión: FC< -2 o FC> 2 (FC, fold change).*

Se llevó a cabo un análisis de enriquecimiento de un conjunto de genes (GSEA, del inglés *Gene Set Enrichment Analysis*) y se vio que la disminución de la expresión de *Myd88* en las células B precursoras preleucémicas *Pax5*$^{+/-}$ aumentaba la expresión de genes implicados en la activación de efectores de señalización (por ejemplo, Notch, de la apoptosis y la señalización IL2-STAT5) asociados a la proliferación y a la supervivencia celular (**Figura 32**). Los genes implicados en la señalización de TNFα (factor de necrosis tumoral α) a través de NFκB, así como en la respuesta de interferón-α e interferón-γ también aumentaron su expresión en las células B preleucémicas *Pax5*$^{+/-}$ (**Figura 32**).

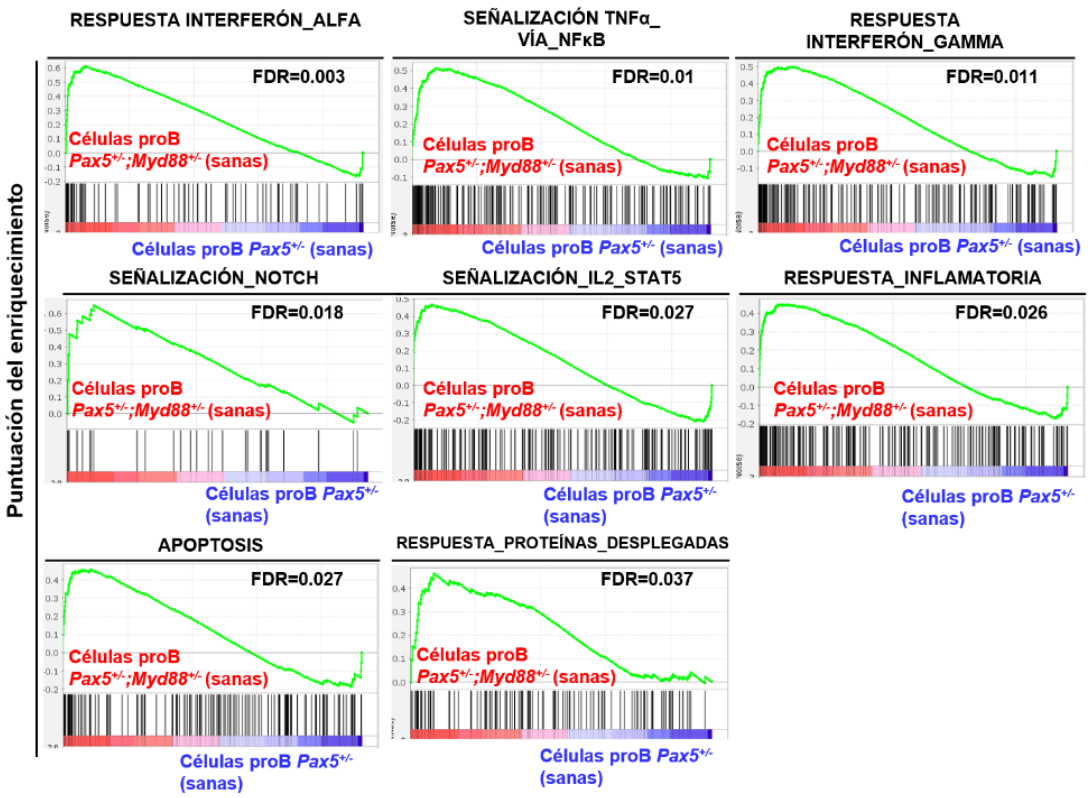



***Figura 32. Análisis del enriquecimiento de la expresión de genes implicados en distintas vías de señalización.*** *Se representan los gráficos GSEAs que muestran que los 8 conjuntos de genes que participan en la vía de respuesta del interferón-α; en la vía señalización de TNFα a través de NFκB; en la respuesta del interferón-γ; en la vía de señalización de Notch; en la vía de señalización IL2-STAT5; en la vía de respuesta inflamatoria; en la vía de apoptosis; y en la respuesta a proteínas desplegadas, están significativamente enriquecidos con un p-valor nominal < 1% en las células proB de ratones Pax5$^{+/-}$;Myd88$^{+/-}$ sanos. El análisis de enriquecimiento se realizó utilizando la colección de sellos distintivos de "Molecular Signatures Database" (MSigDB)*[123].

También se observó que las células proB *Pax5$^{+/-}$;Myd88$^{+/-}$* estaban enriquecidas en genes específicos de los estadios de desarrollo de células B foliculares y células preB murinas (**Figura 33A**). Curiosamente, la disminución de la expresión de *Myd88* en las células proB de ratones *Pax5$^{+/-}$* exacerbó el número de genes expresados diferencialmente que están regulados por *Pax5$^{+/-}$*, como *Cd19*, *Slamf6*, *Bst1*, *Foxo1* o *Bcar3* (**Figura 33B**).

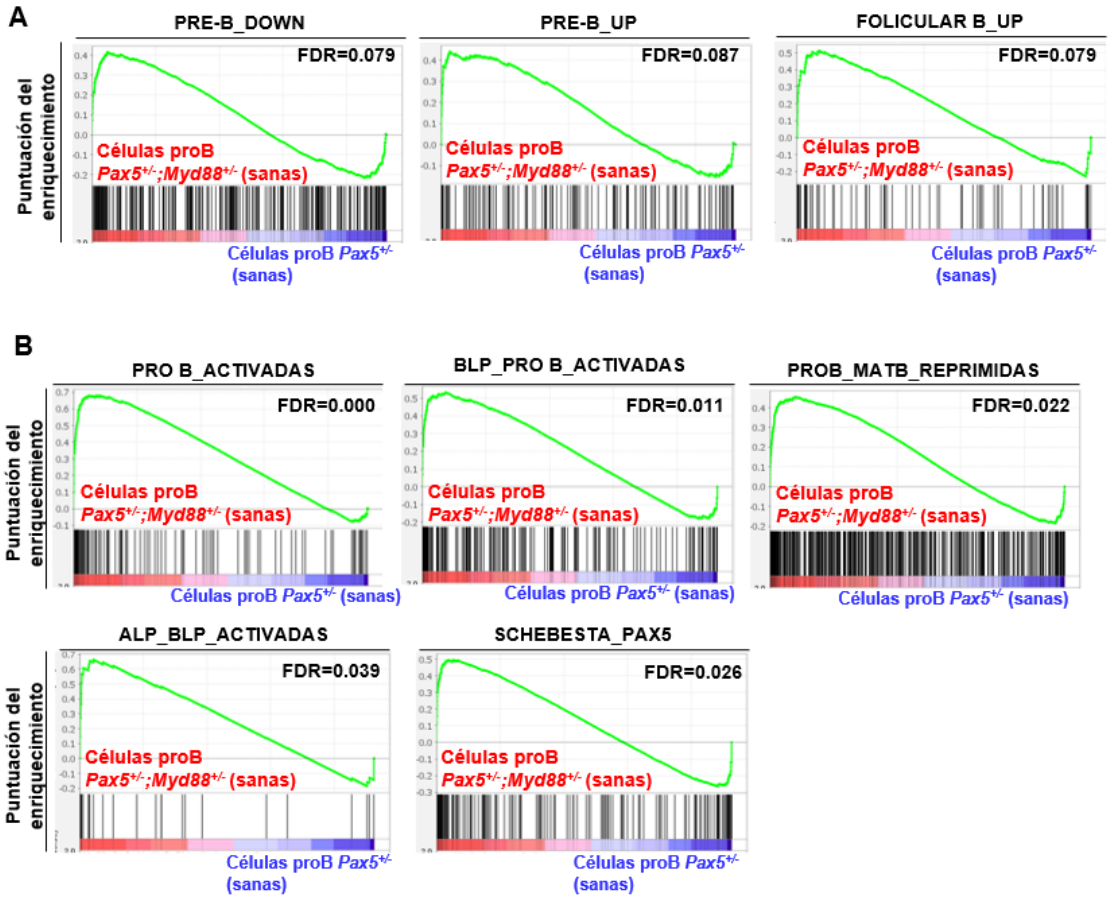



***Figura 33. Impacto de la disminución de la expresión de Myd88 en la expresión génica de las células proB Pax5$^{+/-}$ sanas. A)*** *Se representan los gráficos GSEAs que muestran los estadios de desarrollo de las células B murinas enriquecidos en las células proB de ratones sanos Pax5$^{+/-}$;Myd88$^{+/-}$. Se evaluaron 16 estadios de desarrollo diferentes y se extrajeron conjuntos de genes de Green et al[124]. **B)** GSEAs que muestran que los genes regulados por Pax5 están más enriquecidos en células proB de ratones Pax5$^{+/-}$;Myd88$^{+/-}$ sanos que en células proB de ratones Pax5$^{+/-}$ sanos. Los conjuntos de genes se extrajeron de Revilla et al[125]. y Schebesta et al[126]. (BLP: progenitores linfoides con predisposición a células B; ALP: progenitores totalmente linfoides; MATB: células B maduras). Los conjuntos de genes mostrados están significativamente enriquecidos con un p-valor < 1%.*

Es importante destacar que este perfil transcripcional inflamatorio no está presente con la pérdida completa de *Myd88* en las células proB *Pax5$^{+/-}$*, estando por tanto asociado a niveles intermedios, pero no nulos, de expresión de *Myd88* (**Figura 34**). Estos resultados sugieren que la señalización dependiente de *Myd88* está hiperactivada en las células *Pax5$^{+/-}$*, debido a que dichas células presentan defectos a la hora de sobreexpresar a los inhibidores de la vía de señalización *Myd88*-NFκB (**Figura 32**). Esto está en concordancia con estudios previos que muestran que la deficiencia de *Myd88* produce cambios específicos de tejido en la inducción de citoquinas y en la inflamación[127]. Además, estos resultados concuerdan con trabajos previos *in vitro* que revelan que las citoquinas proinflamatorias como IL6, IL-1β, TNFα[128] o la señalización

dependiente de TGF-β (factor de crecimiento transformante beta)[129,130] promueven la transformación maligna de las células B preleucémicas.

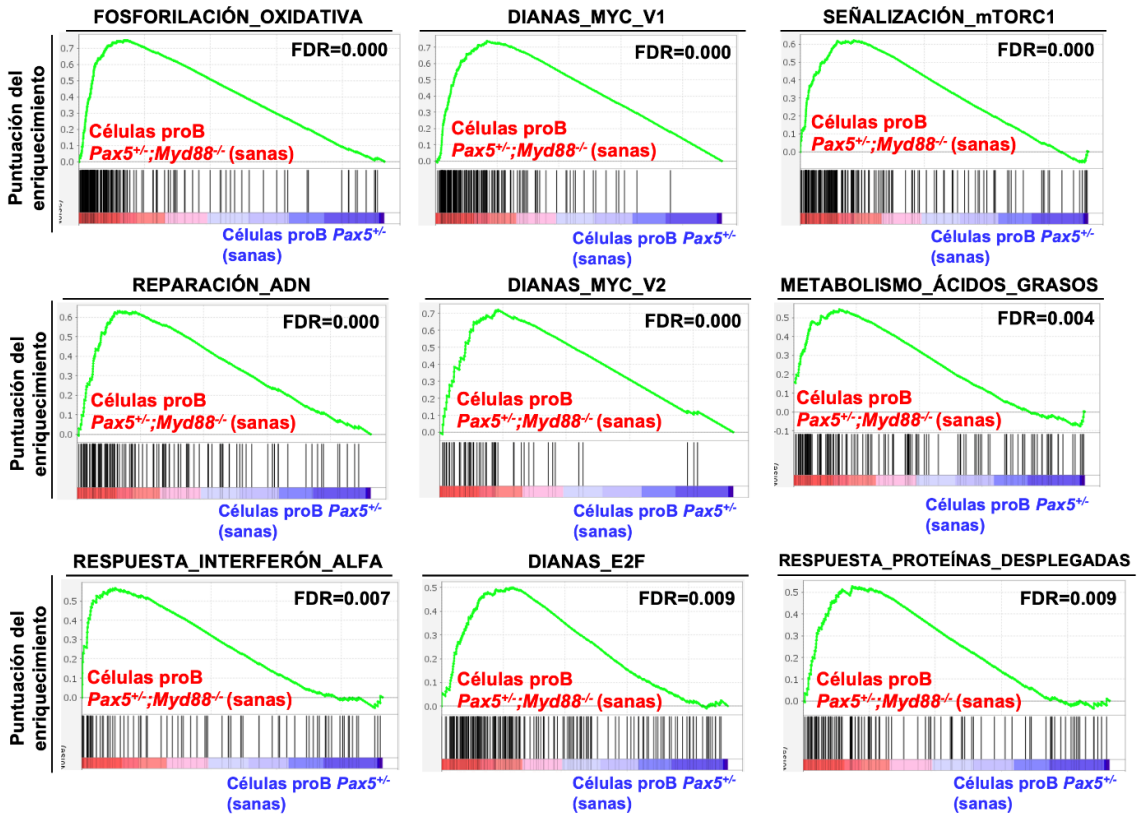



***Figura 34. Impacto de la eliminación de Myd88 en la expresión génica de las células proB de ratones Pax5$^{+/-}$ sanos.*** *Gráficos GSEA que muestran los 9 conjuntos de genes que están significativamente enriquecidos con un p-valor < 1% en las células proB de ratones Pax5$^{+/-}$;Myd88$^{-/-}$ sanos. El análisis de enriquecimiento se realizó utilizando la colección de hallmarks de "Molecular Signatures Database" (MSigDB)*[131].

# DISCUSIÓN

# DISCUSIÓN

## 1. El estrés inmunológico no tiene un papel promotor de la leucemia vinculada a la predisposición genética del gen *ETV6*.

Desde hace algunos años se sabe que la interacción entre genes y ambiente está ligada al desarrollo de la LLA-B infantil[25,26]. Las alteraciones genéticas que conducen a la aparición de un clon preleucémico son múltiples y algunas de ellas ya han sido descritas, como las alteraciones estructurales en el gen *PAX5* y la traslocación *ETV6-RUNX1*[25,26]. Además, estudios recientes han descrito mutaciones germinales en el gen *ETV6* en familias que presentan predisposición a la LLA[78,79]. Sólo el 25-30% de los portadores de las mutaciones en dicho gen terminan desarrollando la leucemia, por lo que, basándonos una vez más en la teoría de los dos eventos, son necesarias alteraciones somáticas secundarias las que den lugar a la enfermedad. De modo que es posible que el desarrollo de la LLA-B vinculada a la predisposición genética asociada al gen *ETV6* esté mediado por un estrés inmunológico como ocurre con otros genes de susceptibilidad (*PAX5* o *ETV6-RUNX1*).

En este trabajo hemos explorado la posibilidad de que el estrés inmunológico promueva el desarrollo de la LLA-B en individuos predispuestos portadores de la mutación p.R355X en el gen *ETV6*. Para ello, hemos usado un modelo de ratón que mimetiza un escenario de predisposición a la leucemia como el que se da en humanos. Nuestros resultados evidencian que la exposición a la infección en ratones *Etv6*$^{R355X}$ genéticamente predispuestos a la leucemia no tiene un efecto promotor de dicha enfermedad. A pesar de haberse encontrado diferencias en cuanto a los porcentajes de células B y células T presentes en la sangre periférica de estos ratones respecto a los ratones WT, estas variaciones no participan o no están directamente relacionadas con la progresión de la enfermedad. De modo que, en el caso de la susceptibilidad genética asociada a *ETV6*, el inmuno-estrés que supone la exposición a patógenos comunes no

consigue promover mutaciones secundarias que permitan la conversión del clon preleucémico.

Los resultados relacionados con el estudio del microbioma en los ratones *Etv6*$^{R355X}$ indican que un tratamiento transitorio con antibióticos sí consigue el efecto deseado de la depleción bacteriana en el intestino de estos animales. Sin embargo, no se puede considerar este inmuno-estrés como responsable de la aparición de la posible leucemia, ya que estos mismos ratones *Etv6*$^{R355X}$ también presentan patologías similares en ausencia de los antibióticos, ambos en condiciones SPF. No obstante, debemos insistir en que la aparición de estas patologías está pendiente de catalogar una vez se disponga de los datos de secuenciación de las muestras patológicas. Por lo tanto, de forma contraria a lo previamente observado con los ratones *Pax5*$^{+/-}$ (donde la alteración del microbioma era suficiente para promover la LLA-B, incluso en ausencia de estímulos infecciosos, hecho que no ocurría sin los antibióticos[100]), en el caso de *ETV6*, un estrés inmunológico como es el uso de antibióticos no es un factor diferencial a la hora de desencadenar la leucemia en ratones con la predisposición genética *Etv6*$^{R355X}$.

Contrariamente a lo que siempre se ha postulado en el campo de la LLA-B infantil, nuestros resultados ponen de manifiesto la existencia de una relación específica entre las alteraciones genéticas y el inmuno-estrés que puede desencadenar la transformación maligna, ya que existe una gran variabilidad en este sentido en la que cada estrés inmunológico puede afectar o no a cada predisposición genética. De este modo, el estudio de la compleja interacción que se da entre los subtipos genéticos y los factores ambientales que impulsan el desarrollo de la LLA-B infantil, puede ayudar a la identificación de dianas terapéuticas para reducir el impacto del estrés inmunitario en portadores predispuestos.

Como ya se ha mencionado, para catalogar los casos de leucemia de los ratones *Etv6*$^{R355X}$ es necesario llevar a cabo la identificación de los eventos secundarios que se adquieren somáticamente y que conducen al desarrollo de la LLA. Para ello se requiere secuenciar el genoma completo de las muestras de ADN tumoral y compararlos con sus correspondientes muestras de ADN control obtenidas de la línea germinal de la cola de dichos ratones. Este es un proceso largo por lo que actualmente no se dispone de los

resultados de la secuenciación, pero estamos a la espera de ellos para confirmar así la aparición de la leucemia en los ratones *Etv6*$^{R355X}$. Todo parece sugerir que sí son casos de leucemia, ya que tanto la edad de aparición de ésta (latencia larga entre los 15 y los 22 meses) como el fenotipo de las células tumorales (B220$^{bajo}$ CD19$^{+/bajo}$ IgM-) llevan a pensar que estas leucemias de los ratones *Etv6*$^{R355X}$ son similares a las que presentan los ratones con la fusión *ETV6-RUNX1*.

Es por todo esto por lo que se puede concluir que los estreses inmunológicos ensayados en este trabajo no tienen un papel promotor de la leucemia vinculada a la predisposición genética del gen *ETV6*. Asimismo, cuando los ratones *Etv6*$^{R355X}$ son expuestos a radiaciones de baja intensidad (2 Gys), tampoco generan LLA-B, aunque este mismo estrés inmunológico la promueve en ratones *Pax5*$^{+/-}$ y *Sca1-ETV6-RUNX1*. Esto lleva a pensar que deben existir otros factores que sean capaces de promover las mutaciones secundarias que conducen al desarrollo de una leucemia completa en la susceptibilidad conferida por alteraciones en el gen *ETV6*.

Además, sería interesante estudiar si pudiera darse el caso de que el modelo de susceptibilidad genética *Etv6*$^{R355X}$ siguiese un modelo de tres eventos en el que además de la propia mutación puntual en el gen *ETV6*, fuera necesaria la cooperación con otra mutación genética. En este sentido, estudios previos han publicado que las LLA-B diploides asociadas a *ETV6* presentan la deleción de *PAX5* con una frecuencia del 30%[80]. Dado que se conoce que los ratones *Pax5*$^{+/-}$ desarrollan LLA-B en condiciones de exposición a la infección con un porcentaje próximo al 22%, sería necesario explorar la incidencia de leucemia en un modelo murino *Etv6*$^{R355X}$;*Pax5*$^{+/-}$ con exposición a patógenos comunes. De esta manera, se esperaría que la cooperación de ambas mutaciones junto a la presencia de un inmuno-estrés como es la infección, aumentara la incidencia de la leucemia en este nuevo modelo de estudio de predisposición genética a la enfermedad.

Estos hallazgos ponen de manifiesto la importancia de descifrar los mecanismos moleculares que tienen lugar cuando la respuesta inmunitaria se desregula tras la exposición a factores ambientales provocando la aparición de la leucemia. El hecho de no conocer estas vías moleculares imposibilita el diseño de un plan de ataque para

impedir la transformación maligna del clon preleucémico. Los nuevos objetivos que se plantean para un futuro próximo pasan por explorar otras vías frecuentemente alteradas, comprender nuevos mecanismos de evasión inmune, así como tratar de trasladar todo este conocimiento a la clínica.

## 2. El estrés inmune suprime la señalización de la inmunidad innata en las células B precursoras preleucémicas para causar leucemia.

Desde hace varias décadas, los datos epidemiológicos han respaldado la asociación entre la incidencia de la LLA-B y los distintos factores de estrés; sin embargo, hasta hace poco, no existían modelos experimentales capaces de proporcionar una explicación biológica adecuada para estos hallazgos[25,26,100,86,112]. En este trabajo se han utilizado modelos animales para aportar pruebas que demuestran que la disbiosis del microbioma intestinal es una característica universal de todas las condiciones genéticas preleucémicas analizadas. Sin embargo, la alteración del microbioma intestinal mediante la depleción bacteriana tiene exclusivamente un impacto específico en el desarrollo de la LLA-B en ratones *Pax5*$^{+/-}$ genéticamente predispuestos. Esto pone de nuevo en relieve la compleja interacción entre los subtipos genéticos y los factores ambientales que impulsan el desarrollo de la LLA-B infantil. De hecho, el estudio de la variación en la susceptibilidad a la LLA-B por subtipos genéticos ha permitido identificar por primera vez una heterogeneidad etiológica en la progresión de la enfermedad entre los subtipos de LLA-B infantil, como también hemos visto en el estudio de los ratones *Etv6*$^{R355X}$.

Asimismo, los resultados presentados en este trabajo sugieren que el uso de la microbiota intestinal como diana terapéutica modificable para reducir o abolir el impacto del estrés inmunitario en portadores predispuestos, podría aplicarse únicamente a algunos subtipos genéticos específicos de LLA-B infantil. En concordancia con este hallazgo, estos estudios muestran que, en ratones *Pax5*$^{+/-}$, el entrenamiento

inmune por exposición temprana a la infección no tiene un efecto protector contra la conversión del clon preleucémico en LLA-B. Este hallazgo entra en conflicto con la hipótesis de "infección retardada" para este subtipo molecular específico de LLA-B, y en su lugar, apoya la teoría de Kinlen de "mezcla de poblaciones". De esta manera, los resultados aquí presentados refuerzan la noción de la variación etiológica en el desarrollo de la LLA-B infantil y la existencia de una interacción específica "gen-ambiente" que desencadena el desarrollo de la LLA-B.

Inesperadamente, los datos mostrados también indican que el desarrollo de la LLA-B entra en la categoría de cánceres asociados con una desregulación en los mecanismos moleculares de la inmunidad innata, la cual desempeña un papel clave en la evolución clonal de los precursores de células B premalignos hacia la LLA-B como resultado del estrés inmunitario. La caracterización molecular de los reguladores de la inmunidad innata que promueven la iniciación de la LLA-B ha permitido la identificación de *Myd88* como un factor clave en este proceso. El aumento de la incidencia de la LLA-B tras la reducción de la expresión de *Myd88* en el modelo murino *Pax5*$^{+/-}$ de predisposición a la leucemia evidenció su papel dependiente de las células B preleucémicas (**Figura 35**). Se demostró que la reducción de expresión de *Myd88* en dichas células incrementaba la expresión de genes implicados en la respuesta al interferón-α, la señalización del TNFα a través de NFκβ, la respuesta al interferón-γ y la activación de efectores de señalización (por ejemplo, Notch, apoptosis y señalización IL2-STAT5) asociados a la proliferación y supervivencia celular. Este papel dependiente de la inflamación de *Myd88* en el desencadenamiento de la LLA-B en células B preleucémicas también se ha demostrado para otros TLRs en el cáncer[88]. La caracterización de la señalización de la inmunidad innata dentro de las células B precursoras preleucémicas ha demostrado que no todas las células preleucémicas son idénticas (**Figura 35**). Según lo propuesto, esta variabilidad podría proporcionar un medio para que las células preleucémicas adapten su comportamiento a las condiciones ambientales y aprovechen las fluctuaciones potenciales en la competencia inmunitaria del organismo en condiciones de estrés[10]. Se especula que dicha adaptación a las nuevas condiciones podría, a su vez, conducir a una reprogramación de las propias células preleucémicas que podría afectar a las moléculas

inmunitarias (como los TLRs y otras moléculas inmunomoduladoras), favoreciendo la adquisición de segundos eventos.

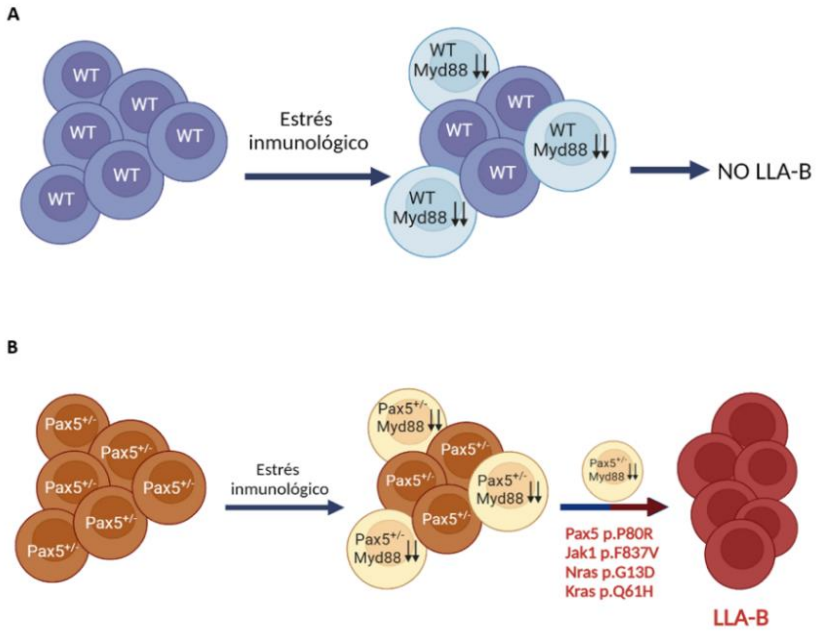



***Figura 35. La supresión de la señalización inmunitaria innata por el estrés inmunitario desencadena el desarrollo de la leucemia a partir de precursores preleucémicos Pax5$^{+/-}$. A)*** *El estrés inmunitario (por ejemplo, exposición a infecciones, tratamiento con antibióticos, etc.) puede afectar tanto a portadores sanos como a portadores preleucémicos Pax5$^{+/-}$ por igual, y en ambos puede provocar una reducción parcial en la expresión de Myd88 en algunos progenitores de células B.* ***B)*** *Sin embargo, sólo en los progenitores preleucémicos Pax5$^{+/-}$ esta disminución de expresión interfiere con la señalización inmune innata, de forma que permite la aparición de mutaciones secundarias que conducen al desarrollo de la leucemia. La incidencia de LLA-B es directamente proporcional al porcentaje de células preleucémicas Pax5$^{+/-}$ que presentan una reducción en la expresión de Myd88.*

Los pasos iniciales en el desarrollo de la LLA-B en niños suelen pasar desapercibidos y, cuando la enfermedad se hace evidente en el momento del diagnóstico, la delimitación de los acontecimientos genéticos consecutivos que conducen progresivamente a la leucemia se complica debido a la acumulación de mutaciones oncogénicas impulsoras y pasajeras. Por lo tanto, la clave para identificar el papel iniciador de la desregulación de la inmunidad innata es el estudio de las células leucémicas en el momento en el que están apareciendo, con el fin de comprender los eventos sucesivos que inician la LLA-B impulsada por la infección. Este enfoque ha permitido identificar mecanismos moleculares previamente desconocidos que tienen lugar en las etapas aún desapercibidas de la formación de la leucemia de células B humanas, abriendo así nuevas oportunidades para el desarrollo de estrategias eficaces

para la detección precoz y la prevención de la LLA-B infantil. En conjunto, estos datos plantean la posibilidad de que el refuerzo de la inmunidad innata pueda representar una estrategia viable para la reducción o prevención del riesgo de sufrir LLA-B infantil. Los resultados presentados en este trabajo también sugieren que la eliminación de estas células preleucémicas reprogramadas, o la protección contra esta adaptación ambiental, podría ayudar a reducir o prevenir el desarrollo leucémico en portadores genéticamente predispuestos de forma más eficaz que las estrategias basadas en el entrenamiento inmunitario general. Del mismo modo, el entrenamiento inmunológico mediante la exposición temprana a la infección no previene a los portadores preleucémicos *Pax5*$^{+/-}$ frente a la LLA-B. En general, estos hallazgos tienen importantes implicaciones para la comprensión y el potencial objetivo terapéutico del estado preleucémico en niños.

# CONCLUSIONES

# CONCLUSIONES

1. El estrés inmunológico causado por la exposición a patógenos comunes en los ratones *Etv6*$^{R355X}$ genéticamente predispuestos no consigue promover la conversión del clon preleucémico. Del mismo modo, la alteración del microbioma intestinal como resultado de un tratamiento antibiótico tampoco supone un efecto directo en la génesis de la LLA asociada a este modelo de ratón. En conjunto, estos resultados sugieren que existe una heterogeneidad etiológica con respecto al papel del estrés inmunológico en el desarrollo de los diferentes subtipos genéticos de LLA-B infantil, contrariamente al pensamiento establecido en el campo sobre una asociación mecánica subyacente común entre la infección y la LLA-B infantil.

2. El entrenamiento inmune mediante la exposición temprana a las infecciones no tiene un efecto protector contra la conversión del clon preleucémico en LLA-B en los ratones *Pax5*$^{+/-}$. Esto sugiere un defecto celular autónomo dentro de las células preleucémicas *Pax5*$^{+/-}$ y refuerza aún más la visión de la variación etiológica en el desarrollo de la leucemia entre los diferentes subtipos moleculares de LLA-B.

3. El inmuno-estrés disminuye la expresión de *Myd88* en las células B preleucémicas *Pax5*$^{+/-}$ tanto humanas como de ratón. Esta supresión de la señalización de la inmunidad innata provoca un aumento en la incidencia de la LLA-B en estos ratones genéticamente predispuestos. Estos hallazgos demuestran que el desarrollo de la LLA-B cae en la categoría de cánceres asociados con una desregulación de los mecanismos moleculares de la inmunidad innata.

4. En conjunto, los hallazgos obtenidos en este proyecto doctoral tienen implicaciones importantes para la comprensión y la posible orientación terapéutica del estado preleucémico en los niños.

# MATERIAL Y MÉTODOS

# MATERIAL Y MÉTODOS

## 1. Modelos animales utilizados en el estudio.

En este estudio se han utilizado varios modelos de ratón descritos ya previamente: *Pax5*$^{+/-}$[132], *Sca1-ETV6-RUNX1*[26], *Sca1-Lmo2*[105] y *Sca1-BCR-ABLp190*[106], y finalmente el modelo de ratón que expresa una mutación puntual en el gen *ETV6* (*Etv6*$^{R355X}$)[80].

### a- Modelo de ratón *Pax5*$^{+/-}$:

El modelo de ratón *Pax5*$^{+/-}$ fue cedido por el laboratorio de M. Busslinger y su generación[132] y caracterización[25] han sido descritas previamente.

### b- Modelo de ratón *Sca1-ETV6-RUNX1*:

El modelo de ratón *Sca1-ETV6-RUNX1* ya ha sido previamente descrito[26] . En estos ratones, la fusión génica *ETV6-RUNX1* se encuentra expresada bajo el control del promotor Sca1, específico de células madre hematopoyéticas. De esta manera, la expresión de este gen de fusión se encuentra restringida a ese compartimento.

### c- Modelo de ratón *Sca1-Lmo2*:

El modelo de ratón *Sca1-Lmo2* fue descrito previamente[105]. En este caso, la expresión del gen *Lmo2* se encuentra restringida a las células madre hematopoyéticas, las cuales están bajo el control del promotor Sca1.

### d- Modelo de ratón *Sca1-BCR-ABLp190*:

El modelo de ratón *Sca1-BCR-ABLp190* también ha sido previamente descrito[106]. En estos ratones, la fusión génica *BCR-ABL* se encuentra expresada bajo el control del promotor Sca1, específico de células madre hematopoyéticas. De esta manera, la expresión de este gen de fusión se encuentra restringida a ese compartimento.

**e- Modelo de ratón *Pax5$^{+/-}$;Myd88$^{+/-}$*:**

Los ratones *Pax5$^{+/-}$* se cruzaron con ratones *Myd88$^{+/-}$* (B6.129P2(SJL)-*Myd88*<tm1.1Defr>/J)[133] para generar ratones *Pax5$^{+/-}$;Myd88$^{+/-}$*.

**f- Modelo de ratón *ETV6$^{R355X}$*:**

Estos ratones presentan la mutación puntual p.Arg355X en el gen *ETV6*, la cual fue descrita previamente[80].

Todos los ratones utilizados en este trabajo tienen un fondo genético mixto C57BL/6J× CBA.

Todos los experimentos con ratones se realizaron siguiendo la normativa legal española y europea aplicable, y han sido previamente autorizados 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 obtuve la capacitación para experimentación en ciencias del animal de laboratorio, categorías a, b y c, por parte de la Consejería de Agricultura y Ganadería de la Junta de Castilla y León.

Los animales fueron generados en un animalario libre de patógenos (Animalario OMG del Centro de Investigación del Cáncer) y mantenidos en el mismo hasta que cumplieron un mes de edad. En ese momento, se trasladaron a un animalario convencional para su exposición a agentes infecciosos naturales como se ha descrito previamente[25]. Las condiciones del animalario incluían una temperatura de 21°C ± 2°C, humedad del 55% ± 10% y un ciclo de 12 horas de luz y 12 horas de oscuridad. El acceso a la comida y la bebida fue *ad libitum*.

Los animales se revisaron y monitorizaron diariamente, y cuando se detectaron signos de leucemia, bien clínicamente (letargia, disminución de la ingesta, pérdida de peso o inflamación de los nódulos linfáticos) o bien por la aparición de células leucémicas en los análisis rutinarios de sangre periférica, el animal fue sacrificado y se realizaron los procedimientos estándar de necropsia. Los órganos principales fueron examinados en el microscopio y muestras representativas de los tejidos se cortaron y

fijaron. Los análisis de supervivencia fueron realizados con la prueba estadística log-rank (Mantel-Cox) utilizando el programa GraphPad Prism v.8.2.1.

El tamaño de las cohortes fue determinado en base a la literatura que describe la incidencia de leucemia por exposición a infección[25,26,86,32,112] y justificada mediante el cálculo de una estimación con una fiabilidad del 90% para poder detectar cambios en la incidencia de la leucemia.

En todos los experimentos se han utilizado ratones de ambos sexos, excepto para los análisis de microbioma, en los que solo se usaron animales hembra para evitar el sesgo por razón de sexo.

## 2. Genotipado de ratones.

El genotipado de los animales se realizó mediante PCR del ADN total extraído de un fragmento de cola de los animales.

### 2.1. Genotipado de los ratones *Pax5*$^{+/-}$.

Para cada ratón se llevaron a cabo dos PCRs en paralelo: una para el alelo WT y otra para el alelo nulo. Se preparó una mezcla de reacción que contenía 2,5µl del oligonucleótido *Pax5*-1 10µM; 2,5 µl del oligonucleótido *Pax5*-2 10µM o 2,5µl del oligonucleótido *Pax5*-3 10µM; 2,5µl del tampón comercial (Ref.2900162, 5PRIME); 2,5µl de dNTPs 2mM; 0,2µl de la enzima Taq ADN polimerasa (Ref. 2900163, 5PRIME); 13,8µl de agua millipore (MPW, del inglés *millipore water*) y 1µl de ADN. Se usaron los siguientes oligos:

Pax5-1: (5'-ACAGTCCCCTTACCCCTTCCCCCAAATGAA-3')

Pax5-2: (5'-GTCCCTCCTAACAAAGTCTCCCTATTCCAC-3')

Pax5-3: (5'-GAGCGAGTAACAACCCGTCGGATTCTCC-3')

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) con las siguientes condiciones: 94°C 4 min; (94°C 1 min; 64°C 1 min; 72°C 2 min) x 40 ciclos; 72°C 7 min. El producto de la amplificación del alelo *WT* tenía un tamaño de 693 pares de bases (pb) y el del alelo nulo de 783 pb. Los tamaños se discriminaron mediante una electroforesis en gel de agarosa al 2%.

## 2.2. Genotipado de los ratones *Sca1-ETV6-RUNX1.*

Los ratones *Sca1-ETV6-RUNX1* se genotiparon mediante PCR de la fusión *ETV6-RUNX1*. Para la PCR de *ETV6-RUNX1* se preparó una mezcla de reacción conteniendo 2,5µL del oligo ETV6-RUNX1-F 10µM; 2,5µL del oligo ETV6-RUNX1-B 10µM; 2,5µL de un buffer comercial (Ref. 2900162, 5PRIME); 2,5µL dNTPs 2mM; 0,2µL del enzima Taq ADN polimerasa (Ref. 2900163, 5PRIME); 13,8µL MPW y 1µL de ADN. Se usaron los siguientes oligos:

ETV6-RUNX1-F: (5'-GCCCCATGGAGAATAATCACTG-3')

ETV6-RUNX1-B: (5'-GCGTCTCTAGAAGGATTCATTC-3')

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) con las siguientes condiciones: 94°C 10min; (94°C 1min; 54°C 1min; 72°C 1min) x 40 ciclos; 72°C 10min. El producto de la amplificación tenía un tamaño de 372 pb. El tamaño se discriminó mediante una electroforesis en gel de agarosa al 2%.

## 2.3. Genotipado de los ratones *Sca1-Lmo2.*

Para la PCR de *Sca1-Lmo2* se preparó una mezcla de reacción conteniendo 2,5µL del oligo Lmo2-F 10µM; 2,5µL del oligo Lmo2-B 10µM; 2,5µL de un buffer comercial (Ref. 2900162, 5PRIME); 2,5µL dNTPs 2mM; 0,2µL del enzima Taq ADN polimerasa (Ref. 2900163, 5PRIME); 13,8µL MPW y 1µL de ADN. Se usaron los siguientes oligos:

Lmo2-F: (5'-CCATCGAAAGGAAGAGCCTG-3')

Lmo2-B: (5'-CTCTCCTGCACAATTTCCGT-3')

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) con las siguientes condiciones: 94°C 10min; (94°C 1min; 54°C 1min; 72°C 1min) x 35 ciclos; 72°C 10min. El producto de la amplificación tenía un tamaño de 237 pb. El tamaño se discriminó mediante una electroforesis en gel de agarosa al 2%.

## 2.4. Genotipado de los ratones *Sca1-BCR-ABLp190*.

Para la PCR de *Sca1-BCR-ABLp190* se preparó una mezcla de reacción conteniendo 2,5µL del oligo P190-F 10µM; 2,5µL del oligo LP190-B 10µM; 2,5µL de un buffer comercial (Ref. 2900162, 5PRIME); 2,5µL dNTPs 2mM; 0,2µL del enzima Taq ADN polimerasa (Ref. 2900163, 5PRIME); 13,8µL MPW y 1µL de ADN. Se usaron los siguientes oligos:

p190-F: (5'-GAACTCGCAACAGTCCTTC-3')

p190-B: (5'-CCATTGTGATTATAGCCTAAG-3')

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) con las siguientes condiciones: 94°C 10min; (94°C 1min; 56°C 1min; 72°C 1min) x 35 ciclos; 72°C 10min. El producto de la amplificación tenía un tamaño de 366 pb. El tamaño se discriminó mediante una electroforesis en gel de agarosa al 2%.

## 2.5. Genotipado de los ratones *Myd88*.

Para la PCR de *Myd88* se preparó una mezcla de reacción conteniendo 2,5µL del oligo 9335 10µM; 2,5µL del oligo oIMR9481 10µM; 2,5µL del oligo oIMR9482 10µM; 2,5µL de un buffer comercial (Ref. 2900162, 5PRIME); 2,5µL dNTPs 2mM; 0,2µL del enzima Taq ADN polimerasa (Ref. 2900163, 5PRIME); 11,3µL MPW y 1µL de ADN. Se usaron los siguientes oligos:

9335: (5’-CCACCCTTGATGACCCCCTA-3’)

oIMR9481: (5’-GTTGTGTGTGTCCGACCGT-3’)

oIMR9482: (5’-GTCAGAAACAACCACCACCATGC-3’)

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) con las siguientes condiciones: 94°C 2min; (94°C 20seg; 65°C 15 seg; 68°C 10 seg) x 10 ciclos; (94°C 15 seg; 60°C 15 seg; 72°C 10 seg) x 28 ciclos; 72°C 2min. El producto de la amplificación en el caso de los WT tenía un tamaño de 266 pb; para los mutantes el tamaño era de 520 pb y para los heterocigotos el tamaño de las bandas es de 520 pb y 266 pb. El tamaño se discriminó mediante una electroforesis en gel de agarosa al 2%.

## 2.6. Genotipado de los ratones *ETV6*$^{R355X}$.

Para la PCR de *ETV6*$^{R355X}$ se preparó una mezcla de reacción conteniendo 2,5µL del oligo ETV6-F 10µM; 2,5µL del oligo ETV6-B 10µM; 2,5µL de un buffer comercial (Ref. 2900162, 5PRIME); 2,5µL dNTPs 2mM; 0,2µL del enzima Taq ADN polimerasa (Ref. 2900163, 5PRIME); 13,8µL MPW y 1µL de ADN. Se usaron los siguientes oligos:

ETV6-F: (5'-AGGATGGGATGAATCGGGAA-3')

ETV6-B: (5'-CCGGGCTCCTTCCTGATAAT-3')

La mezcla de reacción se amplificó usando un termociclador de gradiente (Mastercycler, Eppendorf) con las siguientes condiciones: 94°C 5min; (94°C 1min; 58°C 1min; 72°C 1min) x 25 ciclos; 72°C 10min. El producto de la amplificación tenía un tamaño de 355 pb. El tamaño se discriminó mediante una electroforesis en gel de agarosa al 2%.

Una vez se ha obtenido el fragmento amplificado se utiliza el kit de extracción rápida de gel PureLink (Ref. K210025 - Invitrogen) siguiendo las instrucciones del fabricante, el cual permite purificar de manera rápida y eficaz fragmentos de ADN a partir de geles de agarosa.

Finalmente, se llevó a cabo la secuenciación por Sanger para identificar la mutación deseada en las muestras de ADN mediante un analizador genético 3130 (Applied Biosystems).

## 3. Tratamiento con antibióticos.

Los ratones fueron tratados con un régimen antibiótico según lo descrito[134]. Se les administró durante un período de 8 semanas un cóctel de antibióticos diluidos en el agua de bebida de los animales. Los fármacos utilizados fueron los siguientes: ampicilina, 1 g/L (Ratiopharm); vancomicina, 500 mg/L (Cell Pharm); ciprofloxacina, 200 mg/L (Bayer Vital); imipenem, 250 mg/L (MSD); metronidazol, 1 g/L (Fresenius).

## 4. Análisis celular mediante citometría de flujo.

Se obtuvieron células nucleadas de sangre periférica, timo, bazo, nódulos linfáticos y médula ósea total de ratón mediante lavado con tampón fosfato salino (PBS, de sus siglas en inglés *phosphate-buffered saline*) de los huesos largos. Los eritrocitos se lisaron mediante una incubación a 4°C con un tampón de lisis (RCLB, de las siglas en inglés *red cell lysis buffer*; 0,155 M $NH_4CL$; 10mM $KHCO_3$; 10 mM EDTA pH=7,4) y las células restantes se lavaron con PBS suplementado al 1% con suero fetal bovino (FCS, de las siglas en inglés *fetal calf serum*). Tras la tinción con anticuerpos (**Tabla 2**), las células se lavaron de nuevo con PBS 1% de FCS y se resuspendieron en PBS 1% de FCS que contenía 10 µg/mL de yoduro de propidio (PI, de las siglas en inglés *propidium iodide*) para excluir las células muertas en los análisis.

***Tabla 2 - Relación de anticuerpos utilizados para los análisis por citometría de flujo, concentración utilizada y casa comercial de la que se obtuvieron.***

| Anticuerpo | Dilución empleada | Casa comercial |
|---|---|---|
| anti-B220 (RA3-6B2) | 1:100 | Biolegend |
| anti-CD4 (RM4-5) | 1:250 | Biolegend |

| | | |
|---|---|---|
| anti-CD8a (53-6.7) | 1:250 | Biolegend |
| anti-CD11b/Mac1 (M1/70) | 1:200 | Biolegend |
| anti-CD19 (1D3) | 1:100 | Biolegend |
| anti-CD117/c-Kit (2B8) | 1:200 | Biolegend |
| anti-Ly-6G/Gr1 (RB6-8C5) | 1:100 | Biolegend |
| anti-IgM (R6-60.2) | 1:100 | Biolegend |
| anti-CD25 (PC61) | 1:500 | BD Biosciences |
| anti-CD45 | 1:100 | BD Biosciences |
| anti-Ter119 (TER119) | 1:100 | Biolegend |
| anti-CD16/CD32 (Fc-Block) | 1:100 | BD Biosciences |

Las muestras pertenecientes al apartado de resultados "El estrés inmunológico no tiene un papel promotor de la leucemia vinculada a la predisposición genética del gen ETV6" fueron adquiridas en un citómetro espectral (CYTEK NL-2000) y analizadas mediante el software FCS express. Las muestras incluidas en el apartado de resultados "El estrés inmune suprime la señalización de la inmunidad innata en las células B precursoras preleucémicas para causar leucemia" se adquirieron en un citómetro de flujo AccuriC6 (BD Biosciences) y se analizaron utilizando el software FlowJo.

La fluorescencia específica para los fluorocromos FITC, PE, PI 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 las poblaciones. En cada análisis se examinaron al menos 50.000 células viables (PI-). Para evitar la unión inespecífica de los anticuerpos, las células se incubaron con anti-CD16/CD32 Fc-Block (BD Biosciencies) antes de ser teñidas.

Las poblaciones se identificaron de la siguiente manera: células B maduras en la periferia ($B220^{+}$ $IgM^{+/-}$), células T CD4 ($CD4^{+}$ $CD8^{-}$), células T CD8 ($CD4^{-}$ $CD8^{+}$), granulocitos ($Gr1^{alto}$ $Mac^{alto}$), células T dobles positivas ($CD4^{+}$ $CD8^{+}$), células pro-B y pre-B en la médula ósea ($B220^{bajo}$ $IgM^{-}$), células pro-B en la médula ósea ($CD19^{+}$ c-$Kit^{+}$), células pre-B en la médula ósea ($B220^{bajo}$ $CD25^{+}$ $IgM^{-}$), células B recirculantes en médula

ósea (B220$^{alto}$ IgM$^{+/-}$), células B inmaduras en médula ósea (B220$^{bajo}$ IgM$^{+}$) y células B maduras en bazo y nódulos linfáticos (B220$^{+}$ CD19$^{+}$ y B220$^{+}$ IgM$^{+/-}$).

## 5. Análisis histológicos.

Los animales fueron sacrificados mediante dislocación cervical y las muestras de tejido se fijaron en formaldehído al 4% y se embebieron en parafina. La evaluación patológica se realizó sobre secciones teñidas con hematoxilina-eosina bajo la supervisión del Dr. Óscar Blanco, patólogo experto del Hospital Universitario de Salamanca.

## 6. Análisis de clonalidad de inmunoglobulinas (recombinación VDJ).

Los reordenamientos de las cadenas pesadas de las inmunoglobulinas se amplificaron mediante PCR utilizando distintas parejas de oligonucleótidos (**Tabla 3**). Las condiciones utilizadas en el termociclador consistieron en una activación térmica inicial a 95ºC seguida de 31-37 ciclos de desnaturalización durante 1 minuto a 95ºC, anillamiento durante 1 minuto a 65ºC y elongación durante 1 minuto 45 segundos a 72ºC. A esto le siguió una elongación final durante 10 min a 72ºC.

***Tabla 3. Oligonucleótidos utilizados para la PCR del reordenamiento de las cadenas pesadas de las inmunoglobulinas.*** *R se refiere a purina (A o G). M se refiere a base amina (A o C). Y se refiere a base pirimidina (T o C). W se refiere a bases que crean enlaces débiles (A o T). S se refiere a bases que crean enlaces fuertes (C o G).*

| NOMBRE | DIRECCIÓN | SECUENCIA |
|---|---|---|
| $V_H$J558 | sentido | 5'-CGAGCTCTCCARCACAGCCTWCATGCARCTCARC-3' |
| | antisentido | 5'-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3' |
| $V_H$7183 | sentido | 5'-CGGTACCAAGAASAMCCTGTWCCTGCAAATGASC-3' |
| | antisentido | 5'-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3' |
| $V_H$Q52 | sentido | 5'-CGGTACCAGACTGARCATCASCAAGGACAAYTCC-3' |

| | | |
|---|---|---|
| | antisentido | 5'-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3' |
| $V_H$Gam3.8 | sentido | 5'-CAAGGGACGGTTTGCCTTCTCTTTGGAA-3' |
| | antisentido | 5'-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3' |
| $V_H$3609 | sentido | 5'-KCYYTGAAGAGCCRRCTCACAATCTCC-3' |
| | antisentido | 5'-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3' |
| DH | sentido | 5'-TTCAAAGCACAATGCCTGGCT-3' |
| | antisentido | 5'-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3' |
| Cμ | sentido | 5'-TGGCCATGGGCTGCCTAGCCCGGGACTT-3' |
| | antisentido | 5'-GCCTGACTGAGCTCACACAAGGAGGA-3' |

## 7. Extracción de ADN.

Los tejidos fueron disgregados con un tampón de lisis con proteinasa K (0,25M de sacarosa, 50mM Tris pH 7,4, 25mM KCl, 5mM MgCl2, 1% SDS, 20mM EDTA pH 8 y 1µg de proteinasa K 40µg/mL (Roche)) en volumen variable según la celularidad de cada tejido (entre 500µl y 5ml), durante 16 horas a 55°C. Tras la incubación, se procedió a la extracción del ADN mediante el método fenol-cloroformo. Se agregó un volumen de solución de fenol (Gibco BRL) a las muestras, se homogeneizó y se centrifugó 5 minutos a 14000 rpm, transfiriendo la fase superior a un tubo con un volumen de cloroformo (Merck). De nuevo, la solución se homogeneizó y centrifugó 5 minutos a 14000 rpm, transfiriendo la fase superior a un nuevo tubo con dos volúmenes de etanol absoluto. Seguidamente, la solución se mezcló y se centrifugó en las mismas condiciones. El precipitado formado por el ADN se lavó con un volumen de etanol al 70%, se dejó secar a temperatura ambiente durante 10 minutos y se resuspendió en agua bidestilada. La concentración de ADN se cuantificó mediante la medida de la absorbancia a 260nm con el espectrofotómetro Nanodrop-1000 (THERMO Scientific). Las muestras extraídas se conservaron a -20º C.

## 8. Extracción de ARN.

Los tejidos se disgregaron haciéndolos pasar por filtros de 70µm de nylon (BD FALCON). Posteriormente, se procedió a la extracción de ARN utilizando el kit TRIzol (Life Technologies), de acuerdo con las instrucciones del fabricante. Los restos de fenol procedentes del uso del reactivo TRIzol fueron eliminados con el kit comercial *RNeasy Mini Kit* (Ref. 74104, Qiagen) para evitar posibles interferencias en aplicaciones posteriores.

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).

## 9. Extracción del ADN de heces de ratones.

Se recogieron de manera individual las muestras fecales de cada ratón en diferentes tiempos a lo largo de su vida. El ADN fecal se aisló utilizando el kit *QIAmp ADN Stool Mini Kit* (Qiagen) siguiendo el protocolo de detección de patógenos en heces. El ADN se extrajo siguiendo las instrucciones del fabricante con dos modificaciones: el paso de calentamiento con el tampón ASL se realizó a 95°C y la elución se realizó en 100 µl de tampón AE.

## 10. Secuenciación y análisis del ARN ribosómico 16S.

Se recogieron muestras fecales frescas de cada ratón (*WT* o *Pax5*$^{+/-}$) en puntos temporales definidos. El ADN se extrajo siguiendo las directrices mencionadas anteriormente en el apartado 9. La cuantificación del ADN se realizó por el método PicoGreen (Invitrogen, cat.#P7589) utilizando fluorometría *Victor 3*. Las muestras que pasaron el control de calidad fueron utilizadas para la construcción de bibliotecas.

Para la secuenciación de la región variable V4 del ARNr 16S se siguieron una serie de protocolos que se mencionan a continuación. La selección y amplificación de la región V3-V4 del gen del ARNr 16S se llevó a cabo mediante PCR utilizando *barcoded primers* (oligonucleótidos con un código identificativo). La construcción de la biblioteca fue realizada por Macrogen (https://www.macrogen-europe.com/) utilizando el kit de biblioteca *Herculase II Fusion DNA Polymerase Nextera XT Index Kit V2*. Para verificar el tamaño de los fragmentos amplificados por la PCR, se comprobó la distribución del tamaño con un bioanalizador *Agilent Technologies 2100* utilizando un chip ADN 1000. La secuenciación del extremo emparejado V3-V4 se llevó a cabo utilizando un Illumina MiSeq (La Jolla, CA), de acuerdo con los protocolos del fabricante.

Para el procesamiento de los datos de secuenciación se utilizaron técnicas bioinformáticas como el paquete de análisis del microbioma QIIME 2 2019.10[135]. Los datos de las secuencias sin procesar, en bruto, se de-multiplexaron y se filtraron por calidad con el complemento *q2-demux*, seguido de la eliminación de ruido con DADA2[136] (a través de *q2-dada2*). Todas las variantes de secuencia de los amplicones (ASVs, del inglés *amplicon sequence variants*) se alinearon con mafft[137] (mediante *q2-alignment*) y se utilizaron para construir una filogenia con fasttree2[138] (mediante *q2-phylogeny*). Se estimaron parámetros de diversidad alfa (características observadas y Diversidad Filogenética de Faith[139]), parámetros de diversidad beta (UniFrac ponderado[140], UniFrac no ponderado[141], distancia de Jaccard y diferencia de Bray-Curtis[142]) y análisis de coordenadas principales (PCoA, del inglés *Principle Coordinate Analysis*) utilizando *q2-diversity* después de que las muestras fueran "rarificadas" o submuestreadas sin reemplazo a 80992 secuencias por muestra. La taxonomía se asignó a las ASV utilizando el clasificador de taxonomía *q2-feature-classifier*[143] *classify-sklearn naïve Bayes* contra las secuencias de referencia SILVA 132 99% OTUs[144] . Para el análisis comparativo de la taxonomía de los diferentes grupos, la tabla de características se filtró para mantener las características con una frecuencia mínima de 10 y presentes en, al menos, 10 muestras. El análisis comparativo se determinó utilizando el *Linear Discriminant Analysis Effect Size*[145].

## 11. PCR array de la vía de señalización murina de los TLRs.

Se extrajo ARN de las células precursoras B presentes en la médula ósea (B220$^{bajo}$ IgM$^{-}$) de ratones *WT* y *Pax5*$^{+/-}$ utilizando el reactivo TRIzol (Invitrogen). El ADNc preparado con el kit *RT2 First Strand* (Qiagen) fue preamplificado utilizando el kit de síntesis de ADNc de ciclo celular *RT² HT First Strand Kit* (Qiagen). La expresión de los genes TLR y sus vías de señalización se analizaron en las muestras utilizando el *RT² Profiler PCR Array* (Qiagen) de ratón. Se utilizó la fuente *PCR array DataAnalysisV4web* (Qiagen) para calcular los valores de ciclo umbral (ΔCt) y obtener los valores ΔCt normalizados respecto de los genes housekeeping. Para determinar las diferencias de expresión relativa entre las poblaciones de células B, se realizaron análisis comparativos (2-ΔΔCt) utilizando el software *PCR Array Data Analysis Software* (Qiagen).

## 12. PCR cuantitativa en tiempo real de *Myd88* (q-PCR).

Se analizó mediante PCR cuantitativa (q-PCR) la expresión de *Myd88* tanto en las muestras preleucémicas (células B precursoras (B220$^{bajo}$ IgM$^{-}$) aisladas de la médula ósea de ratones *WT* (n=3), *Pax5*$^{+/-}$(n=3) y *Pax5*$^{+/-}$ tratados con antibióticos durante 8 semanas (n=3)), como en muestras leucémicas (células leucémicas de ratones *Pax5*$^{+/-}$ tratados con antibióticos durante 8 semanas (n=3) y células leucémicas de ratones *Myd88*$^{+/-}$*;Pax5*$^{+/-}$ (n=3)).

El ADNc para uso en estudios de PCR cuantitativa se sintetizó utilizando la transcriptasa reversa *SuperScript VILOTM cDNA Synthesis Kit* (Invitrogen). Las reacciones de PCR en tiempo real se realizaron en una máquina Eppendorf *MasterCycler Realplex*. Los ensayos utilizados para la PCR cuantitativa están disponibles comercialmente en IDT (*Integrated DNA Technologies*): *Myd88* (ID del ensayo: Mm.PT.58.8716051.gs) y GAPDH (ID del ensayo: Mm.PT.39a.1). Además, las sondas se diseñaron de manera que no se detectara ADN genómico durante la PCR. La medición de la expresión del producto génico GAPDH se utilizó como control endógeno. La médula ósea total de ratones

*Myd88*$^{-/-}$ se utilizó como control negativo. Todas las muestras se analizaron por triplicado.

Además, se utilizó el método comparativo CT (ΔΔCt) para calcular la expresión relativa del transcrito de interés y un control positivo. El cambio en el ciclo umbral (ΔCt) de cada muestra se calculó como el valor Ct del gen analizado (diana) menos el valor Ct de GAPDH (control endógeno). El ΔΔCt de cada muestra se obtuvo restando el valor ΔCt de referencia del valor ΔCt de la muestra. El valor ΔCt de referencia utilizado fue el ΔCt obtenido del bazo total de ratones *WT* inmunizados. Se comparó el cambio de expresión en cada grupo, calculado como 2-ΔΔCt de la muestra.

## 13. Cuantificación de los niveles de citoquinas en el suero.

Los niveles séricos de citoquinas presentes en los ratones se analizaron mediante el sistema de inmunoensayo *Cytometric Bead Array Inflammation Kit* (BD Biosciences), el cual evalúa simultáneamente IL6, IL10, MCP1, IFNɣ, TNFα e IL12p70. Brevemente explicado, 50 µL de suero se incubaron durante 2 horas a temperatura ambiente con 50 µL de *beads* recubiertas con anticuerpos anticitoquinas (*mAb-coated beads*) y 50 µL de detector de anticuerpos anticitoquinas conjugado con PE (*phycoerythrin*). Tras este período de incubación, las muestras se lavaron una vez (5 minutos a 200 g) para eliminar el exceso de anticuerpos detectores. Inmediatamente después, se realizó la adquisición de datos en un citómetro de flujo FACSCanto II (BD Biosciences) utilizando el software FACSDiva (BD Biosciences). Durante la adquisición se almacenó la información de al menos 3000 eventos positivos para cada *bead* analizada (el número total de *beads* fue superior a 9000). Para el análisis se utilizó el programa *FCAP Array Software* v3.0 (BD Biosciences).

## 14. Cultivo de células pro-B prelucémicas.

Las células pro-B se obtuvieron de la médula ósea de ratones *WT*, *Pax5*$^{+/-}$ y *Pax5*$^{+/-}$*;Myd88*$^{+/-}$ de 2-3 semanas de edad. El aislamiento de las células pro-B fue realizado mediante separación celular activada por magnetismo (MACS, de las siglas en inglés *Magnetic Activated Cell Sorting*) positiva para B220.

Los cultivos de células pro-B se realizaron utilizando el siguiente medio: medio IMDM + 1mM de L-glutamina + HEPES, 2% de FCS inactivado por calor, 1 mM de penicilina-estreptomicina (BioWhittaker), 0,03% de primatone RL (Sigma), 50 µM de β-Mercaptoetanol y 2 µg/ml de ciproxina. Las células obtenidas mediante la separación por MACS se cultivaron sobre células estromales ST2 en el medio descrito anteriormente suplementado con 5 ng/ml de IL7 recombinante murina.

## 15. Análisis de datos mediante *microarrays* de expresión.

El ARN total se aisló en dos pasos, utilizando TRIzol (Life Technologies) seguido de la purificación con *RNeasy Mini-Kit* (Qiagen), siguiendo el protocolo de limpieza del ARN del fabricante con el tratamiento opcional de *On-Column DNasa*. La integridad y la calidad del ARN se verificaron mediante electroforesis en gel de agarosa y su concentración se midió utilizando *Agilent* (Agilent Technologies). Las muestras se analizaron utilizando el software *Applied Biosystems Transcriptome Analysis Consol*e (TAC) v4.0.1 con los siguientes parámetros: (i) *Analysis Version*: version 1; (ii) *Summarization Method*: Gene Level – RMA; (iii) *Genome Version*: mm10 (*Mus musculus*) y (iv) *Annotation*: MoGene-1_0-st-v1.na36.mm10.transcript.csv.

Los genes expresados diferencialmente se evaluaron mediante criterios estadísticos específicos (p-valor <0,05 y cambio de expresión absoluto >2). El análisis de enriquecimiento de la expresión de determinados grupos de genes se llevó a cabo utilizando el software GSEAv4.3.2[131,146,147], la colección de genes hallmark[131], el conjunto de genes de las fases de desarrollo de las células B murinas[124] y el conjunto de genes regulados por *Pax5*[125,126]. Las firmas de expresión genética que están específicamente

sobreexpresadas o subexpresadas en las LLA-B humanas[119,148,149,122] se compararon también con las encontradas en las leucemias murinas de los estudios usando GSEA.

Los datos de expresión mediante microarrays están depositados en la base de datos de *Gene Expression Omnibus* (GEO) del NCBI, disponibles con el número de acceso de la serie GEO GSE221597.

## 16. Análisis de enriquecimiento (GSEA).

El análisis de enriquecimiento de la expresión de determinados grupos de genes se llevó a cabo utilizando el software GSEAv4.3.2[131,146,147], la colección de genes hallmark[131], el conjunto de genes de las fases de desarrollo de las células B murinas[124] y el conjunto de genes regulados por *Pax5*[125,126]. Las firmas de expresión genética que están específicamente sobreexpresadas o subexpresadas en las LLA-B humanas[119-122] se compararon también con las encontradas en las leucemias murinas de los estudios usando GSEA.

## 17. Secuenciación del genoma completo (WGS).

Primeramente, se prepararon librerías de ADN genómico a partir de ADN fragmentado (extraído de muestras de médula ósea o nódulos linfáticos tumorales con más de un 70% de células tumorales y muestras de ADN de la cola como muestras germinales) con el kit *HyperPrep Library Preparation Kit* (Roche PN 07962363001). A continuación, se realizó la secuenciación de 150 ciclos por partes en el NovaSeq 6000 (Illumina) y las lecturas pareadas de Illumina se pre-procesaron y mapearon al genoma de referencia del ratón (mm10) con BWA[150]. Se usó una aproximación en conjunto para buscar mutaciones somáticas (SNV/indels) con varias herramientas publicadas, incluyendo Mutect2[151], SomaticSniper[152], VarScan2[153], MuSe[154] y Strelka[155]. Los resultados que tenían al menos dos fuentes se consideraron como mutaciones fiables.

Estos resultados fueron revisados manualmente para una lectura más profunda y un mapeo de calidad para retirar cualquier artefacto.

Las alteraciones somáticas en el número de copias se determinaron por CNV kit[156]. Para las variantes estructurales SV, se implementaron cuatro fuentes de SV en el flujo de trabajo para la búsqueda de SV, incluyendo Delly[157], Lumpy[158], Manta[159] y Gridss[160]. Los resultados de SV que pasaron los filtros de control de calidad de cada fuente se fusionaron usando SURVIVOR[161] y se genotiparon por SBtyper[162]. Los resultados encontrados se revisaron manualmente usando IGV para apoyar los recuentos más flojos y las lecturas discordantes en ambos extremos de un supuesto sitio de SV.

Los datos de secuenciación del genoma completo están accesibles en la base de datos "*Sequence Read Archive* (SRA)" del NCBI, disponibles con el número de acceso de BioProject PRJNA912234.

## 18. Secuenciación Sanger.

Las mutaciones fueron validadas mediante secuenciación Sanger en un analizador genético 3130 (Applied Biosystems) utilizando los siguientes pares de oligonucleótidos (5' → 3'):

mJak3: sentido: CGGGATGTGGGGCTTTAACT
antisentido: GCAGACACGGGGTATAGTGG

mJak1: sentido: CCAGACAGCCAGGAGAACAG
antisentido: CGTCTGCATAGTACCCACCC

mPax5: sentido: CTCGTACATGCACGGAGACA
antisentido: GGACCCTTCAGTACACCAGC

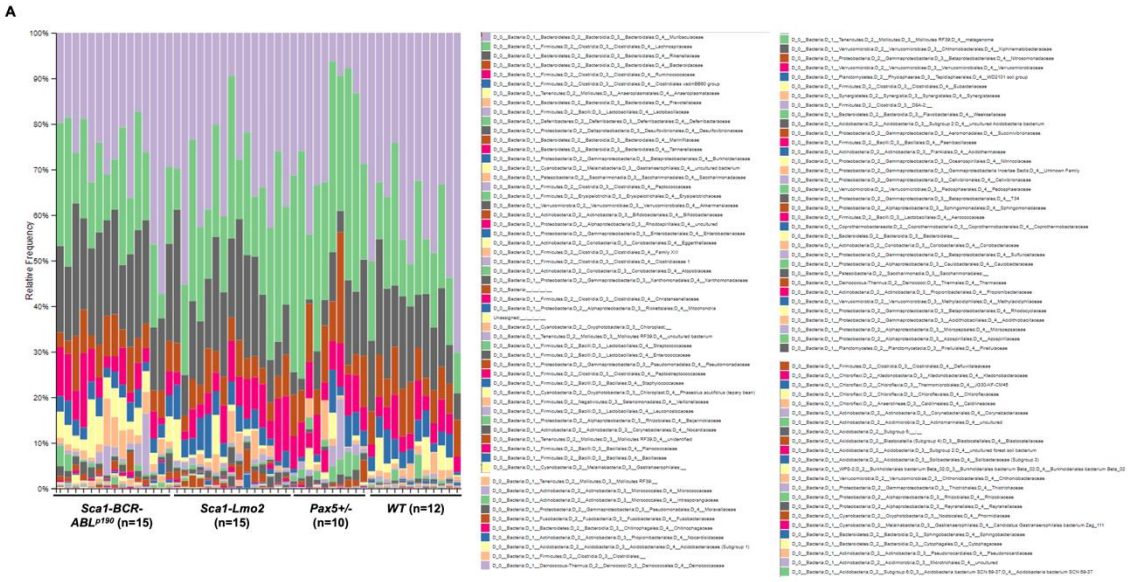
## 19. Análisis estadísticos.

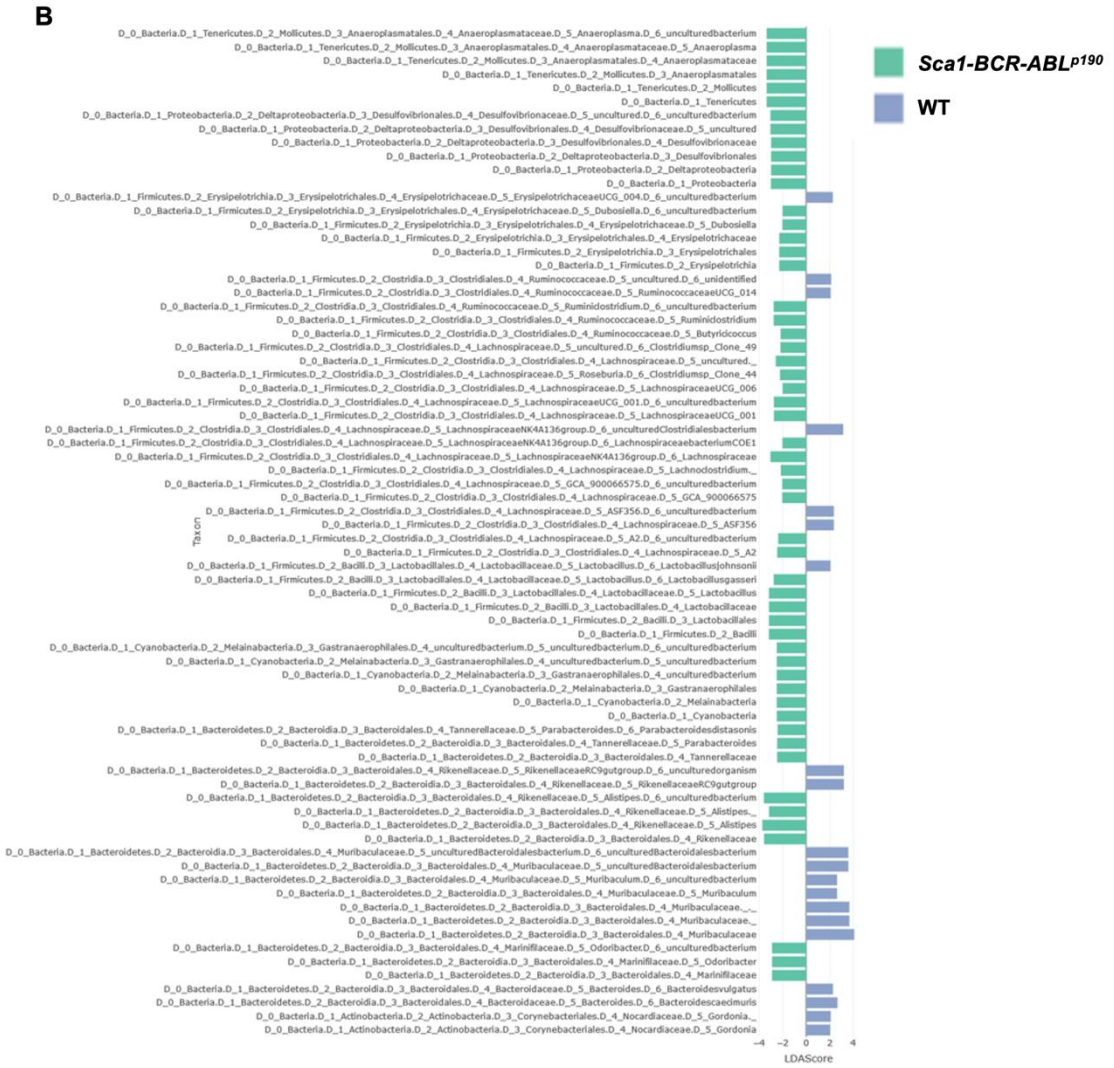
Los análisis estadísticos se realizaron con el programa GraphPad Prism v8.1.2 (GraphPad Software). La significación estadística se calculó mediante el test t no

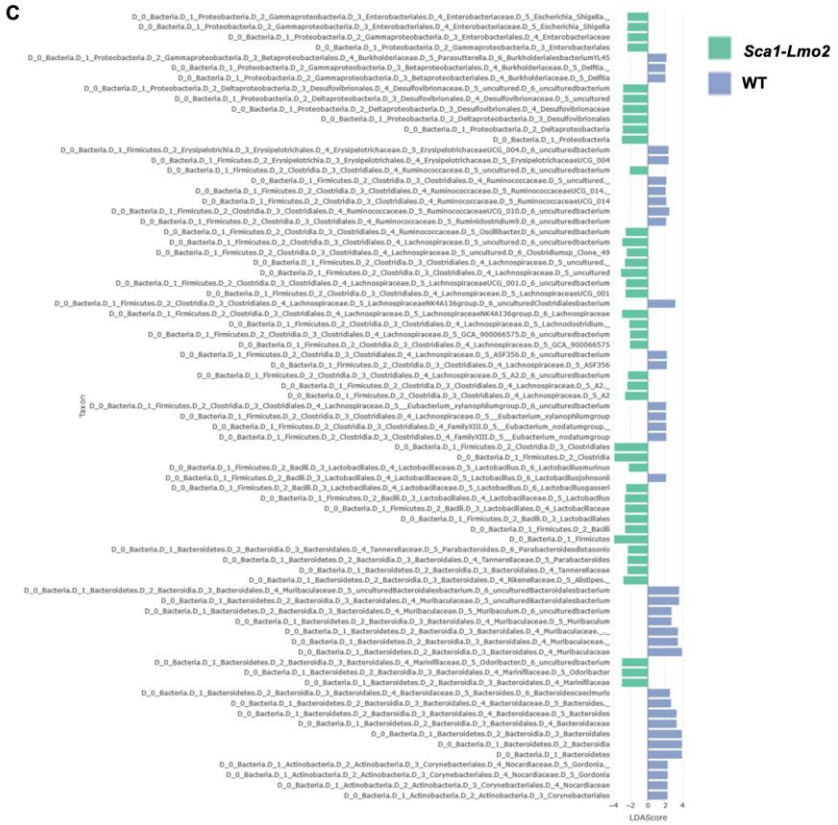
pareado. La normalidad de los datos se comprobó mediante la prueba de D'Agostino y Pearson (nivel alfa = 0,05). Los análisis de supervivencia se realizaron mediante la prueba de Log-rank (Mantel-Cox). Para comparar la incidencia de leucemia se utilizó el test de Fisher. Los p-valores se representan en las figuras sólo en los casos de diferencias significativas.

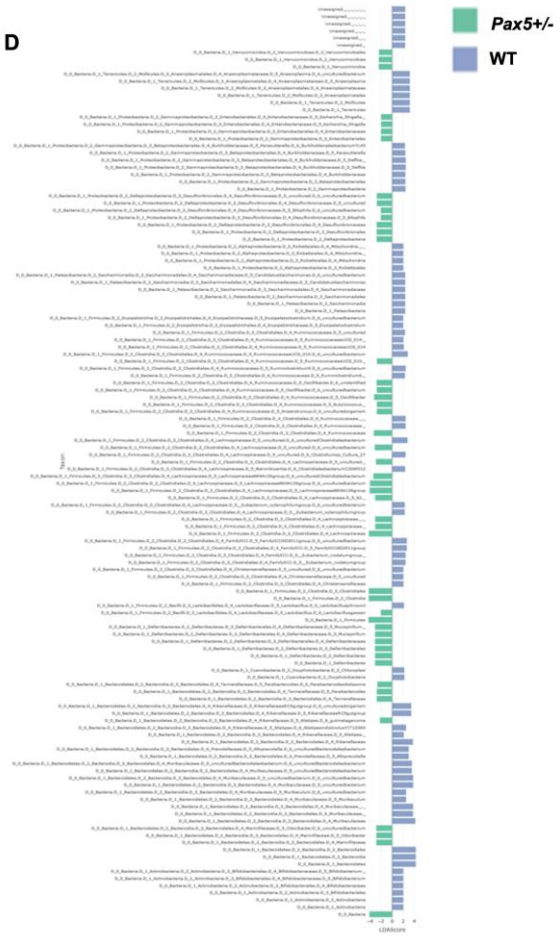
# ANEXOS

# ANEXOS


A
Relative Frequency
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Sca1-BCR-ABLp190 (n=15)
Sca1-Lmo2 (n=15)
Pax5+/- (n=10)
WT (n=12)


B
Sca1-BCR-ABLp190
WT
Taxon
LDAScore
−4
−2
0
2
4

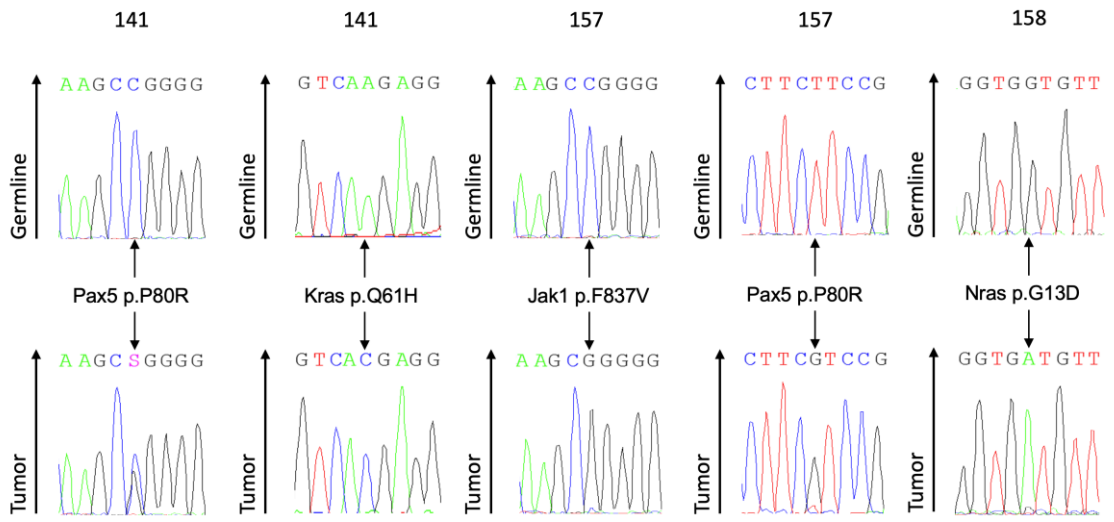

C
Sca1-Lmo2
WT
Taxon
LDAScore


D
Pax5+/-
WT
Taxon
LDAScore

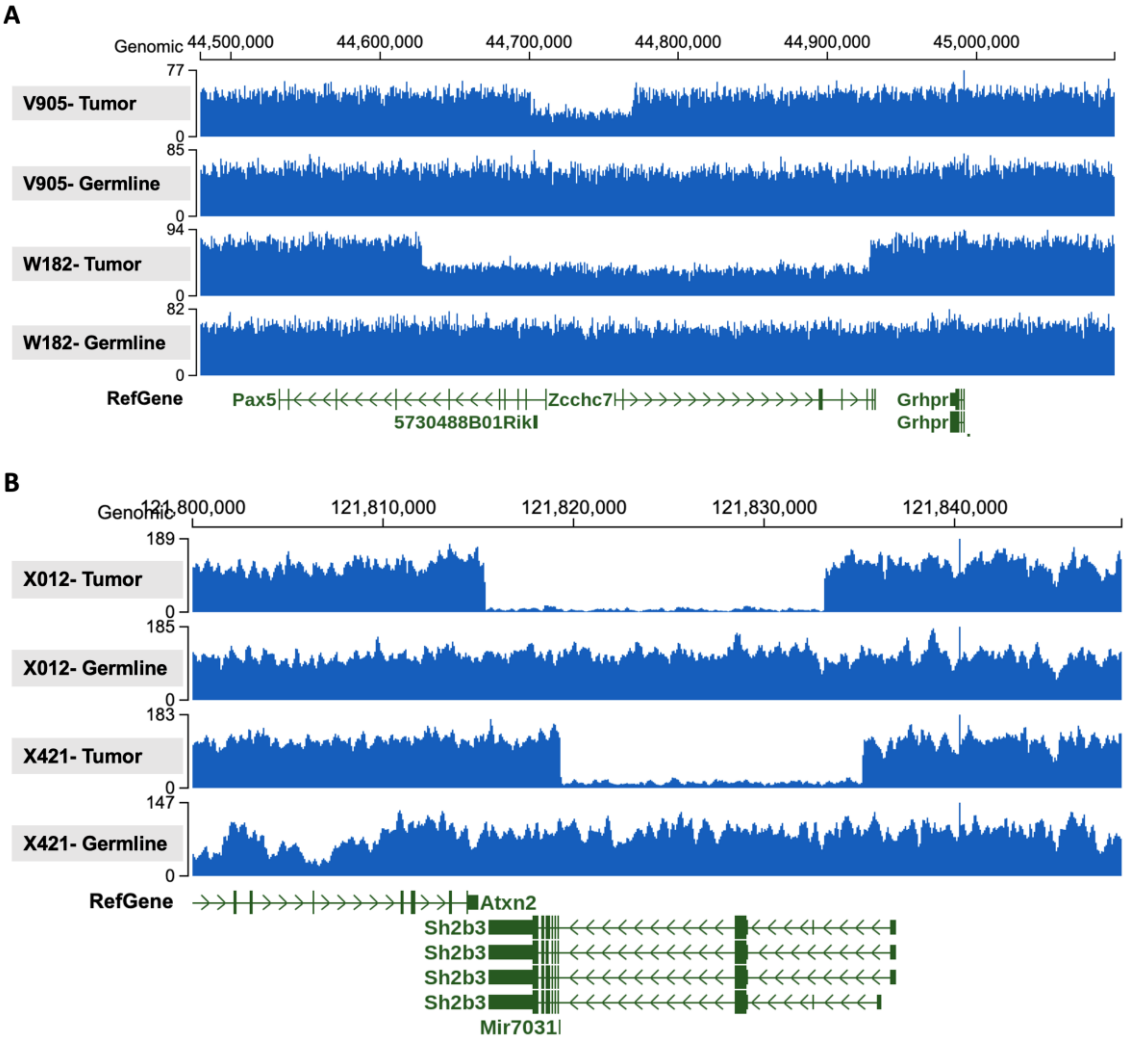
***Anexo 1. Análisis del microbioma intestinal de ratones Sca1-BCR-ABLp190, Sca1-Lmo2, Pax5$^{+/-}$ y WT alojados en un animalario libre de patógenos. A)*** *Diagrama de barras de taxones de las firmas microbianas V4-16S (V4-ASVs) agrupadas por genotipo de ratón (Sca1-BCR-ABLp190 n=15, Sca1-Lmo2 n=15, Pax5$^{+/-}$ n=10 93 y WT n=12) utilizando 5 niveles taxonómicos.* ***B)*** *Los análisis LefSe indican taxones diferencialmente abundantes en ratones Sca1-BCR-ABLp190 frente a ratones WT. Gráfico de análisis diferencial que muestra las 74 unidades taxonómicas con diferencias significativas entre las muestras de ADN fecal de ratones Sca1-BCR-ABLp190 (verde) y ratones WT.* ***C)*** *Los análisis LefSe indican taxones diferencialmente abundantes de ratones Sca1-Lmo2 frente a ratones WT. Gráfico de análisis diferencial que muestra las 77 unidades taxonómicas con diferencias significativas entre muestras de ADN fecal de ratones Sca1-Lmo2 (verde) y ratones WT.* ***D)*** *Los análisis LefSe indican taxones diferencialmente abundantes de ratones Pax5$^{+/-}$ vs WT. Gráfico de análisis diferencial que muestra las 127 unidades taxonómicas con diferencias significativas entre muestras de ADN fecal de ratones Pax5$^{+/-}$ (verde) y ratones WT. La clasificación taxonómica se realizó utilizando la base de datos SILVA para asignar etiquetas taxonómicas a los V4-ASVs, que están anotados en la leyenda. Los análisis LEfSe se realizaron utilizando α=1E-05 y una puntuación LDA mínima = 2. LDA, análisis discriminante lineal.*

***Anexo 2: Informe de Seguimiento Sanitario de la Instalación Convencional.*** *La tabla muestra los patógenos analizados para controlar el estado de salud de los animales alojados en la instalación convencional durante el tiempo en que los animales han sido estudiados.*

| ANIMALARIO CONVENCIONAL – INFORME CONTROL SANITARIO | | |
|---|---|---|
| **AÑO** | **1$^{er}$ año** | **2º año** |
| **VIRUSES** | | |
| MAD (Adenovirus type 1 and 2) | NEG | NEG |
| MCMV (Mouse cytomegalovirus) | NEG | NEG |
| LCMV (Lymph choriomeningitis) | NEG | NEG |
| Ectromelia | NEG | NEG |
| EDIM | NEG | NEG |
| Hantaan virus | NEG | NEG |
| MHV (Mouse hepatitis virus) | POS | POS |
| MVM (Minute virus) | NEG | NEG |
| MPV (Parvovirus type 1 and 2) | NEG | NEG |
| PVM (Pneumonia virus) | NEG | NEG |
| Reovirus type 3 | NEG | NEG |
| Sendai | NEG | NEG |
| TMEV (Theiler´s murine encephalomyel) | NEG | NEG |
| K virus | NEG | NEG |
| Polyoma virus | NEG | NEG |
| MNV (Norovirus) | POS | POS |
| **BACTERIA** | | |
| Bordetella bronchiseptica | NEG | NEG |
| Car-bacillus | NEG | NEG |
| Citrobacter rodentium | NEG | NEG |
| Clostridium piliforme (Tyzzer´ disease) | NEG | NEG |
| Corynebacterium kutscheri | NEG | NEG |
| Mycoplasma pulmonis | NEG | NEG |
| Pasteurella pneumotropica | NEG | NEG |
| Pasteurella multocida | NEG | NEG |
| Salmonella spp | NEG | NEG |
| Streptobacillus moniliformis | NEG | NEG |
| Streptococci beta hemolíticos (A,G) | NEG | NEG |
| Streptococcus pneumoniae | NEG | NEG |
| Helicobacter spp | POS | POS |
| **PARASITES** | | |
| Ectoparasites | NEG | NEG |
| Helminth (A.tetraptrera, Syphacia spp) | POS | POS |
| Cestods (Hymenolepis spp) | NEG | NEG |
| Eimeria spp | NEG | NEG |
| Entamoeba muris | NEG | NEG |
| Giardia spp | NEG | NEG |
| Trichomonas muris | POS | POS |
| Spironucleus spp | NEG | NEG |
| Encephalitozoon cuniculi | NEG | NEG |
| **NECROPSY** | No gross lesion. | No gross lesion. |
| **NEG: Negative** | | |
| **POS: Positive** | | |




***Anexo 3. Secuenciación Sanger de tumores en ratones Pax5$^{+/-}$ expuestos tempranamente a la infección.*** *Todas las variantes encontradas por WES fueron confirmadas por secuenciación Sanger. Del mismo modo que las leucemias Pax5$^{+/-}$ del grupo de exposición tardía, los tumores Pax5$^{+/-}$ de exposición temprana mostraron mutaciones recurrentes que afectaban al alelo Pax5 WT restante: se identificaron las mutaciones Pax5$^{P80R}$, Jak1$^{F837V}$, Nras$^{G13D}$ y Kras$^{Q61H.ç}$*



***Anexo 4. Visualización de mutaciones somáticas en ratones leucémicos Pax5$^{+/-}$;Myd88$^{+/-}$.*** *Visualización del navegador del genoma del ADN tumoral (T) y el ADN de la línea germinal (G) de ratones Pax5$^{+/-}$;Myd88$^{+/-}$ portadores de tumores* ***A)*** *Deleciones somáticas del gen Pax5 identificadas por WGS en ratones leucémicos Pax5$^{+/-}$;Myd88$^{+/-}$ (V905 y W182).* ***B)*** *Deleciones somáticas del gen Sh2b3 identificadas por WGS en ratones leucémicos Pax5$^{+/-}$;Myd88$^{+/-}$ (X012 y X421).* ***C)*** *Validación de la deleción del exón 2 de Pax5 de la línea germinal mediante WGS de todos los ratones Pax5$^{+/-}$;Myd88$^{+/-}$ analizados mediante WGS. En los ratones Pax5+/-, el exón 2 del gen Pax5 (que es indispensable para la unión del ADN) fue reemplazado por los genes de Escherichia coli lac Z y de resistencia a la neomicina.*

# BIBLIOGRAFÍA

# BIBLIOGRAFÍA


1. Malard, F. & Mohty, M. Acute lymphoblastic leukaemia. *Lancet* **395**, 1146-1162 (2020).
2. Kantarjian, H. M., Keating, M. J. & Freireich, E. J. Toward the potential cure of leukemias in the next decade. *Cancer* **124**, 4301-4313 (2018).
3. de Barrios, O. & Parra, M. Epigenetic Control of Infant B Cell Precursor Acute Lymphoblastic Leukemia. *Int. J. Mol. Sci.* **22**, 3127 (2021).
4. Hunger, S. P., Mullighan, C. G. Acute Lymphoblastic Leukemia in Children. *N. Engl. J. Med.* **373**, 1541-1552 (2015).
5. Chang, J. H., Marcus, K. J., Poppe, M. M., Hua, C. & Esiashvili, N. Acute lymphoblastic leukemia. *Pediatr. Blood Cancer* **68**, e28371 (2021).
6. Siegel, R. L., Miller, K. D., Fuchs, H. E., Jemal, A. Cancer statistics, 2022. *CA Cancer J Clin.* **72**, 7-33 (2022).
7. American Cancer Society. Cancer Facts & Figures 2022. (2022).
8. Wiemels, J. Perspectives on the causes of childhood leukemia. *Chem Biol Interact.* **196**, 59-67 (2012).
9. Downing, J. R. *et al.* The pediatric cancer genome project. *Nat. Genet.* **44**, 619-622 (2012).
10. Cobaleda, C., Vicente-Dueñas, C. & Sanchez-Garcia, I. Infectious triggers and novel therapeutic opportunities in childhood B cell leukaemia. *Nat. Rev. Immunol.* **1**, 570-581 (2021).
11. Macarthur, A. C. *et al.* Risk of childhood leukemia associated with vaccination, infection, and medication use in childhood: the Cross-Canada Childhood Leukemia Study. *Am. J. Epidemiol.* **167**, 598-606 (2008).
12. Inaba, H. & Mullighan, C. G. Pediatric acute lymphoblastic leukemia. *Haematologica* **105**, 2524-2539 (2020).
13. Inaba, H., Greaves, M. & Mullighan, C. G. Acute lymphoblastic leukaemia. *Lancet* **381**, 1943-1955 (2013).
14. Kato, M. & Manabe, A. Treatment and biology of pediatric acute lymphoblastic leukemia. *Pediatr. Int.* **60**, 4-12 (2018).
15. Pui, C. H., Nichols, K. E. & Yang, J. J. Somatic and germline genomics in paediatric acute lymphoblastic leukaemia. *Nat. Rev. Clin. Oncol.* **16**, 227-240 (2019).
16. Greaves, M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat. Rev. Cancer* **6**, 193-203 (2006).
17. Doll, R. & Wakeford, R. Risk of childhood cancer from fetal irradiation. *Br. J. Radiol.* **70**, 130-139 (1997).
18. Karipidis, K. K., Benke, G., Sim, M. R., Kauppinen, T. & Giles, G. Occupational exposure to ionizing and non-ionizing radiation and risk of glioma. *Occup Med.* **57**, 518-24 (2007).
19. Smith, M. A., Mccaffrey, R. P. & Karp, J. E. The Secondary Leukemias: Challenges and Research Directions. *J Natl Cancer Inst.* **88**, 407-425 (1996).
20. Caughey, R. W. & Michels, K. B. Birth weight and childhood leukemia: a meta-analysis and review of the current evidence. *Int J Cancer* **124**, 2658-70 (2009).
21. Gruhn, B. *et al.* Prenatal origin of childhood acute lymphoblastic leukemia, association with birth weight and hyperdiploidy. *Leukemia* **22**, 1692-1697 (2008).
22. Mcnally, R. J. Q. & Eden, T. O. B. An infectious aetiology for childhood acute leukaemia:

a review of the evidence. *Br J Haematol.* **127**, 243-63 (2004).

23. Schwab, C. & Harrison, C. J. Acute lymphoblastic leukaemia. *Methods Mol. Biol.* **730**, 99-117 (2011).
24. Descatha, A., Jenabian, A., Conso, F. F. & Ameille, J. Occupational exposures and haematological malignancies: overview on human recent data. *Cancer Causes Control* **16**, 939-53 (2005).
25. Martín-Lorenzo, A. *et al.* Infection exposure is a causal factor in B-cell precursor acute lymphoblastic leukemia as a result of Pax5-inherited susceptibility. *Cancer Discov.* **5**, 1328-1343 (2015).
26. Rodríguez-Hernández, G. *et al.* Infection exposure promotes ETV6-RUNX1 precursor B-cell leukemia via impaired H3K4 demethylases. *Cancer Res.* **77**, 4365-4377 (2017).
27. Kinlen, L. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish new town with nuclear reprocessing sites in Britain. *Lancet* **2**, 1323-7 (1988).
28. Kinlen, L. Childhood leukaemia, nuclear sites, and population mixing. *Br. J. Cancer* **104**, 12-18 (2011).
29. Greaves, M. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat. Rev. Cancer* **18**, 471-484 (2018).
30. Greaves, M. F. Speculations on the cause of childhood acute lymphoblastic leukemia. *Leukemia* **2**, 120-125 (1988).
31. Roman, E. *et al.* Childhood acute lymphoblastic leukemia and infections in the first year of life: A report from the United Kingdom Childhood Cancer Study. *Am. J. Epidemiol.* **165**, 496-504 (2007).
32. Vicente-Duenãs, C. *et al.* An intact gut microbiome protects genetically predisposed mice against leukemia. *Blood* **136**, 2003-2017 (2020).
33. Roy, A. A «gut feeling» about precursor B-ALL. *Blood* **136**, 1995-1996 (2020).
34. Guarner, F. & Malagelada, J. Gut flora in health and disease. *Lancet* **361**, 512-519 (2003).
35. Mcfarland, L. V. Antibiotic-associated diarrhea: epidemiology, trends and treatment. *Futur. Microbiol.* **3**, 563-578 (2008).
36. Mayer, E. A., Tillisch, K. & Gupta, A. Gut/brain axis and the microbiota. *J Clin Invest.* **125**, 926-38 (2015).
37. Zhao, Q. & Elson, C. O. Adaptive immune education by gut microbiota antigens. *Immunology* **154**, 28-37 (2018).
38. Levy, M., Kolodziejczyk, A. A., Thaiss, C. A. & Elinav, E. Dysbiosis and the immune system. *Nat. Rev. Immunol.* **17**, 219-232 (2017).
39. Berg, G. *et al.* Microbiome definition re-visited: old concepts and new challenges. *Microbiome* **8**, 103 (2020).
40. Dethlefsen, L. & Relman, D. A. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci USA* **108 Suppl**, 4554-61 (2011).
41. Lange, K. & Stallmach, A. Effects of Antibiotics on Gut Microbiota. *Dig Dis.* **34**, 260-268 (2016).
42. Mallol, J., Foliaki, S., Pearce, N., Montefort, S. & Zealand, N. Antibiotic use in infancy and symptoms of asthma, rhinoconjunctivitis, and eczema in children 6 and 7 years old: International Study of Asthma and Allergies in Childhood Phase III. *J Allergy Clin Immunol.* **124**, 982-989 (2009).

43. Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci Transl Med.* **8**, 343-382 (2016).
44. Mori, H. *et al.* Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc Natl Acad Sci USA* **99**, 8242-7 (2002).
45. Haven, N. *et al.* Five percent of healthy newborns have an ETV6-RUNX1 fusion as revealed by DNA-based GIPFEL screening. *Blood* **131**, 821-826 (2018).
46. Hein, D., Borkhardt, A. & Fischer, U. Insights into the prenatal origin of childhood acute lymphoblastic leukemia. *Cancer Metastasis Rev.* **39**, 161-171 (2020).
47. Maia, A. T. *et al.* Prenatal origin of hyperdiploid acute lymphoblastic leukemia in identical twins. *Leukemia* **17**, 2202-2206 (2003).
48. Tebbi, C. K. Etiology of Acute Leukemia: A Review. *Cancers (Basel).* **103**, 38-48 (2021).
49. Papaemmanuil, E. *et al.* RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat. Genet.* **46**, 116-125 (2014).
50. Teng, G. *et al.* RAG Represents a Widespread Threat to the Lymphocyte Genome. *Cell* **162**, 751-765 (2015).
51. Mullighan, C. G. *et al.* BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature* **453**, 110-114 (2008).
52. Gröbner, S., Worst, B., Weischenfeldt, J. et al. The landscape of genomic alterations across childhood cancers. *Nature* **555**, 321–327 (2018).
53. Fischer, U. *et al.* Cell Fate Decisions: The Role of Transcription Factors in Early B-cell Development and Leukemia. *Blood Cancer Discov.* **1**, 224-233 (2020).
54. Kimura, S., Mullighan, C. G. & Program, H. M. Molecular markers in ALL: Clinical implications. *Best Pr. Res Clin Haematol* **33**, 1-24 (2020).
55. Raboso-Gallego, J., Casado-García, A., Isidro-Hernández, M. & Vicente-Dueñas, C. Epigenetic Priming in Childhood Acute Lymphoblastic Leukemia. *Front. Cell Dev. Biol.* **7**, (2019).
56. Gu, Z. *et al.* PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat. Genet.* **51**, 296-307 (2019).
57. Nachman, J. B. *et al.* Outcome of treatment in children with hypodiploid acute lymphoblastic leukemia. *Blood* **110**, 1112-1116 (2007).
58. Harrison, C. J. Cytogenetics of paediatric and adolescent acute lymphoblastic leukaemia. *Br J Haematol.* **144**, 147-156 (2009).
59. Moorman, A. V. New and emerging prognostic and predictive genetic biomarkers in B-cell precursor acute lymphoblastic leukemia. *Haematologica* **101**, 407-416 (2016).
60. Pui, B. C. *et al.* 11q23/MLL Rearrangement Confers a Poor Prognosis in infants with acute lymphoblastic leukemia. *J Clin Oncol.* **12**, 909-915 (1994).
61. Schwab, C. & Harrison, C. J. Advances in B-cell precursor acute lymphoblastic leukemia genomics. *HemaSphere* **2**, e53 (2018).
62. Mullighan, C. G. The genomic landscape of acute lymphoblastic leukemia in children and young adults. *Hematol. Am Soc Hematol Educ Progr.* **2014**, 174-180 (2014).
63. Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-764 (2007).
64. Dittmer, J. & Nordheim, A. Ets transcription factors and human disease. *Biochim Biophys Acta* **1377**, F1-11 (1998).

65. Dc, W. *et al.* The 12;21 translocation involving TEL and deletion of the other TEL allele: two frequently associated alterations found in childhood acute lymphoblastic leukemia. *Blood* **87**, 2891-2899 (1996).
66. Oikawa, T. & Yamada, T. Molecular biology of the Ets family of transcription factors. *Gene* **303**, 11-34 (2003).
67. Kim, C. A. *et al.* Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *EMBO J.* **20**, 4173-4182 (2001).
68. Guidez, F. *et al.* Recruitment of the nuclear receptor corepressor N-CoR by the TEL moiety of the childhood leukemia-associated TEL-AML1 oncoprotein. *Blood* **96**, 2557-2561 (2000).
69. Rasighaemi, P., Onnebo, S. M. N., Liongue, C. & Ward, A. C. ETV6 (TEL1) regulates embryonic hematopoiesis in zebrafish. *Haematologica* **100**, 23-31 (2015).
70. Wang, L. C. *et al.* Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. *EMBO J.* **16**, 4374-4383 (1997).
71. Hock, H. *et al.* Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. *Genes Dev.* **18**, 2336-2341 (2004).
72. De Braekeleer, E. *et al.* ETV6 fusion genes in hematological malignancies: A review. *Leuk. Res.* **36**, 945-961 (2012).
73. Rasighaemi P, L. C. and W. A. ETV6 (TEL1) in Blood Cell Development and Malignancy. *J. Blood Disord.* **1**, 7 (2014).
74. Bohlander, S. K. ETV6: A versatile player in leukemogenesis. *Semin Cancer Biol.* **15**, 162-174 (2005).
75. Moriyama, T. *et al.* Germline genetic variation in ETV6 and risk of childhood acute lymphoblastic leukaemia: A systematic genetic study. *Lancet Oncol.* **16**, 1659-1666 (2015).
76. Noetzli, L. *et al.* Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. *Nat. Genet.* **47**, 535-538 (2015).
77. Topka, S. *et al.* Germline ETV6 Mutations Confer Susceptibility to Acute Lymphoblastic Leukemia and Thrombocytopenia. *PLoS Genet.* **11**, 1-14 (2015).
78. Rampersaud, E. *et al.* Germline deletion of ETV6 in familial acute lymphoblastic leukemia. *Blood Adv.* **3**, 1039-1046 (2019).
79. Järviaho, T. *et al.* Predisposition to childhood acute lymphoblastic leukemia caused by a constitutional translocation disrupting ETV6. *Blood Adv.* **3**, 2722-2731 (2019).
80. Nishii, R. *et al.* Molecular basis of ETV6 -mediated predisposition to childhood acute lymphoblastic leukemia. *Blood* **137**, 364-373 (2021).
81. Mackereth, C. D. *et al.* Diversity in structure and function of the Ets family PNT domains. *J. Mol. Biol.* **342**, 1249-1264 (2004).
82. Feurstein, S. & Godley, L. A. Germline ETV6 mutations and predisposition to hematological malignancies. *Int. J. Hematol.* **106**, 189-195 (2017).
83. Cobaleda, C., Schebesta, A., Delogu, A. & Busslinger, M. Pax5: The guardian of B cell identity and function. *Nat. Immunol.* **8**, 463-470 (2007).
84. McManus, S. *et al.* The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. *EMBO J.* **30**, 2388-2404 (2011).
85. Cobaleda, C., Jochum, W. & Busslinger, M. Conversion of mature B cells into T cells by

dedifferentiation to uncommitted progenitors. *Nature* **449**, 473-477 (2007).

86. Rodríguez-Hernández, G. *et al.* Infectious stimuli promote malignant B-cell acute lymphoblastic leukemia in the absence of AID. *Nat. Commun.* **10**, 5563 (2019).
87. Shah, S. *et al.* A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nat. Genet.* **45**, 1226-1231 (2013).
88. Man, S. M. & Jenkins, B. J. Context-dependent functions of pattern recognition receptors in cancer. *Nat. Rev. Cancer* **22**, 397-413 (2022).
89. Rakoff-nahoum, S. & Medzhitov, R. Toll-like receptors and cancer. *Nat. Rev. Cancer* **9**, 57-63 (2009).
90. Goldszmid, R. S. & Trinchieri, G. The price of immunity. *Nat. Genet.* **13**, 932-8 (2012).
91. Chiron, D., Bekeredjian-ding, I. & Pellat-deceunynck, C. Review article Toll-like receptors: lessons to learn from normal and malignant human B cells. *Blood* **112**, 2205-2213 (2008).
92. Hua, Z. & Hou, B. TLR signaling in B-cell development and activation. *Cell. Mol. Immunol.* **10**, 103-106 (2013).
93. Salcedo, R., Cataisson, C., Hasan, U., Yuspa, S. H. & Trinchieri, G. MyD88 and its divergent toll in carcinogenesis. *Trends Immunol.* **34**, 379-389 (2013).
94. Hong, EH., Chang, SY., Lee, BR., et al. Blockade of Myd88 signaling induces antitumor effects by skewing the immunosuppressive function of myeloid-derived suppressor cells. *Int J Cancer* **132**, 2839-48 (2013).
95. Macedo, L. *et al.* Wound Healing Is Impaired in MyD88-Deficient Mice. *Am. J. Pathol.* **171**, 1774-1788 (2007).
96. Rakoff-nahoum, S., Paglino, J., Eslami-varzaneh, F., Edberg, S. & Medzhitov, R. Recognition of Commensal Microflora by Toll-Like Receptors Is Required for Intestinal Homeostasis. *Cell* **118**, 229-241 (2004).
97. Iimuro, Y. *et al.* Role of innate immune response in liver regeneration. *J. Gastroenterol. Hepatol.* **22**, S57-8 (2007).
98. Tian, M. *et al.* B Cell Intrinsic MyD88 Signaling Promotes Initial Cell Proliferation and Differentiation To Enhance the Germinal Center Response to a Virus-like Particle. *J Immunol.* **200**, 937-948 (2018).
99. Liu, X. *et al.* Expression of the prosurvival kinase HCK requires PAX5 and mutated MYD88 signaling in MYD88-driven B-cell lymphomas. *Blood* **4**, (2020).
100. Vicente-Duenãs, C. *et al.* An intact gut microbiome protects genetically predisposed mice against leukemia. *Blood* **136**, (2020).
101. Cobaleda, C., Vicente-Dueñas, C., Ramírez-Orellana, M. & Sanchez-Garcia, I. Revisiting the concept of childhood preleukemia. *Trends in Cancer* **8**, 887-889 (2022).
102. Belkaid, Y. & Hand, T. W. Review Role of the Microbiota in Immunity and Inflammation. *Cell* **157**, 121-141 (2014).
103. Pushalkar, S. *et al.* The Pancreatic Cancer Microbiome Promotes Oncogenesis by Induction of Innate and Adaptive Immune Suppression. *Cancer Discov.* **8**, 403-416 (2018).
104. Greaves, M., Cazzaniga, V. & Ford, A. Can we prevent childhood Leukaemia? *Leukemia* **35**, 1258-1264 (2021).
105. García-Ramírez, I. *et al.* Lmo2 expression defines tumor cell identity during T-cell leukemogenesis. *EMBO J.* **37**, (2018).
106. Martín-Lorenzo, A. *et al.* Loss of Pax5 exploits sca1-BCR-ABLp190 susceptibility to confer the metabolic shift essential for pB-ALL. *Cancer Res.* **78**, 2669-2679 (2018).

107. Rudant, J. *et al.* Systematic Reviews and Meta- and Pooled Analyses Childhood Acute Lymphoblastic Leukemia and Indicators of Early Immune Stimulation : A Childhood Leukemia International Consortium Study. *Am. J. Epidemiol.* **181**, 549-562 (2015).
108. Kamper-Jørgensen M, Woodward A, Wohlfahrt J, Benn CS, Simonsen J, Hjalgrim H, S. K. Childcare in the first 2 years of life reduces the risk of childhood acute lymphoblastic leukemia. *Leukemia* **22**, 189-193 (2008).
109. Ajrouche, R. *et al.* Childhood acute lymphoblastic leukaemia and indicators of early immune stimulation : the Estelle study ( SFCE ). *Br. J. Cancer* **112**, 1017-1026 (2015).
110. Amitay EL, K.-B. L. Breastfeeding and Childhood Leukemia Incidence A Meta-analysis and Systematic Review. *JAMA Pediatr.* **169**, 1-9 (2015).
111. Urayama, K. Y., Buffler, Ã. P. A., Gallagher, E. R., Ayoob, J. M. & Ma, X. A meta-analysis of the association between day-care attendance and childhood acute lymphoblastic leukaemia. *Int J Epidemiol.* **39**, 718-732 (2010).
112. Casado-García, A. *et al.* Transient Inhibition of the JAK/STAT Pathway Prevents B-ALL Development in Genetically Predisposed Mice. *Cancer Res.* **82**, 1098-1109 (2022).
113. Søegaard, S. H., Schmiegelow, K., Rostgaard, K. & Hjalgrim, H. Childcare attendance and risk of childhood acute lymphoblastic leukaemia : A register study based on the Danish childcare database. *Int J Cancer* **152**, 1817-1826 (2023).
114. Auer, F. *et al.* Inherited susceptibility to pre B-ALL caused by germline transmission of PAX5 c.547G>A. *Leukemia* **28**, 1136-1138 (2014).
115. Shah, S. *et al.* A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nat. Genet.* **45**, 1226-1231 (2013).
116. Duployez, N. *et al.* Germline PAX5 mutation predisposes to familial B-cell precursor acute lymphoblastic leukemia. *Blood* **137**, 1424-1428 (2021).
117. Case, M. *et al.* Mutation of Genes Affecting the RAS Pathway Is Common in Childhood Acute Lymphoblastic Leukemia. *Cancer Res.* **68**, 6803-6809 (2008).
118. Kohlmann, A. *et al.* An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics : the Microarray Innovations in LEukemia study prephase. *Br J Haematol.* **142**, 802-807 (2008).
119. Linka, Y. *et al.* The impact of TEL-AML1 (ETV6-RUNX1) expression in precursor B cells and implications for leukaemia using three different genome-wide screening methods. *Blood Cancer J.* **3**, 1-9 (2013).
120. Juric, D. *et al.* Differential Gene Expression Patterns and Interaction Networks in BCR-ABL – Positive and – Negative Adult Acute Lymphoblastic Leukemias. *J Clin Oncol.* **25**, 1341-1349 (2015).
121. Chiaretti, S. *et al.* Gene Expression Profiles of B-lineage Adult Acute Lymphocytic Leukemia Reveal Genetic Patterns that Identify Lineage Derivation and Distinct Mechanisms of Transformation. *Clin. Cancer Res.* **11**, 7209-7219 (2005).
122. Kohlmann, A. *et al.* Pediatric acute lymphoblastic leukemia (ALL) gene expression signatures classify an independent cohort of adult ALL patients. *Leukemia* **18**, 63-71 (2004).
123. Liberzon, A. *et al.* The Molecular Signatures Database. *Cell Syst* **1**, 417-425 (2015).
124. Green, M. R. *et al.* Signatures of murine B-cell development implicate Yy1 as a regulator of the germinal center-specific program. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 2873-2878 (2011).

125. Revilla-I-Domingo, R. *et al.* The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J.* **31**, 3130-3146 (2012).
126. Schebesta, A. *et al.* Transcription Factor Pax5 Activates the Chromatin of Key Genes Involved in B Cell Signaling, Adhesion, Migration, and Immune Function. *Immunity* **27**, 49-63 (2007).
127. Behera, A. K. *et al.* MyD88 Deficiency Results in Tissue-Specific Changes in Cytokine Induction and Inflammation in Interleukin-18-Independent Mice Infected with Borrelia burgdorferi. *Infect Immun* **74**, 1462-1470 (2006).
128. Beneforti, L. *et al.* Pro-inflammatory cytokines favor the emergence of ETV6-RUNX1-positive pre-leukemic cells in a model of mesenchymal niche. *Br J Haematol.* **190**, 262-273 (2020).
129. Ford, A. M. *et al.* The TEL-AML1 leukemia fusion gene dysregulates the TGF- β pathway in early B lineage progenitor cells. *J Clin Invest.* **119**, 826-836 (2009).
130. Portale, F. *et al.* Activin A contributes to the definition of a pro-oncogenic bone marrow microenvironment in t ( 12 ; 21 ) preleukemia. *Exp. Hematol.* **73**, 7-12.e14 (2019).
131. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015).
132. Urbánek, P., Wang, Z. Q., Fetka, I., Wagner, E. F. & Busslinger, M. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5 BSAP. *Cell* **79**, 901-912 (1994).
133. Adachi, O. *et al.* Targeted disruption of the MyD88 gene results in loss of IL-1- and IL- 18-mediated function. *Immunity* **9**, 143-150 (1998).
134. Cording, S. *et al.* Commensal microbiota drive proliferation of conventional and Foxp3 + Regulatory CD4 + T cells in mesenteric lymph nodes and Peyer's patches . *Eur. J. Microbiol. Immunol.* **3**, 1-10 (2013).
135. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852-857 (2019).
136. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581-583 (2016).
137. Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059-3066 (2002).
138. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, (2010).
139. Faith, D. P. Conservation Evaluation and Phylogenetic Diversity. *Biol. Conserv.* **61**, 1-10 (1992).
140. Lozupone, C. A., Hamady, M., Kelley, S. T. & Knight, R. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. *Appl. Environ. Microbiol.* **73**, 1576-1585 (2007).
141. Lozupone, C. & Knight, R. UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**, 8228-8235 (2005).
142. Dudman, W. F. & Belbin, L. Numerical Taxonomic Analysis of Some Strains of Rhizobium spp. That Uses a Qualitative Coding of Immunodiffusion Reactions . *Appl. Environ. Microbiol.* **54**, 1825-1830 (1988).

143. Bokulich, N. A. *et al.* Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 1-17 (2018).
144. McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610-618 (2012).
145. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011).
146. Mootha, V. K. et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267-273 (2003).
147. 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 (2005).
148. Juric, D. *et al.* Differential gene expression patterns and interaction networks in BCR-ABL-positive and -negative adult acute lymphoblastic leukemias. *J. Clin. Oncol.* **25**, 1341-1349 (2007).
149. Chiaretti, S. *et al.* Gene expression profiles of B-lineage adult acute lymphocytic leukemia reveal genetic patterns that identify lineage derivation and distinct mechanisms of transformation. *Clin. Cancer Res.* **11**, 7209-7219 (2005).
150. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).
151. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213-219 (2013).
152. Larson, D. E. *et al.* Somaticsniper: Identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* **28**, 311-317 (2012).
153. Koboldt, D. C. *et al.* VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568-576 (2012).
154. Fan, Y. *et al.* MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. *Genome Biol.* **17**, 178 (2016).
155. Kim, S. *et al.* Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591-594 (2018).
156. Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput. Biol.* **12**, 1-18 (2016).
157. Rausch, T. *et al.* DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, 333-339 (2012).
158. Layer, R. M., Chiang, C., Quinlan, A. R. & Hall, I. M. LUMPY: A probabilistic framework for structural variant discovery. *Genome Biol.* **15**, 1-19 (2014).
159. Chen, X. *et al.* Manta: Rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220-1222 (2016).
160. Cameron, D. L. *et al.* GRIDSS: Sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. *Genome Res.* **27**, 2050-2060 (2017).
161. Jeffares, D. C. *et al.* Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat. Commun.* **8**, 1-11 (2017).
162. Chiang, C. *et al.* SpeedSeq: Ultra-fast personal genome analysis and interpretation. *Nat. Methods* **12**, 966-968 (2015).