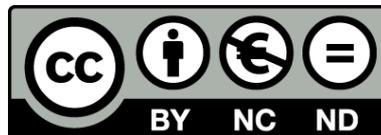




UNIVERSITAT DE  
BARCELONA

## Implicación de las alteraciones genéticas en el pronóstico y progresión de la leucemia linfoblástica aguda de tipo T (LLA-T) del adulto

Celia González Gil



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – SenseObraDerivada 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – SinObraDerivada 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-NoDerivs 4.0. Spain License.**



UNIVERSITAT DE  
BARCELONA



Institut de Recerca  
CONTRA LA LEUCÈMIA  
**Josep Carreras**

Programa de Doctorado en Biomedicina

IMPLICACIÓN DE LAS ALTERACIONES GENÉTICAS  
EN EL PRONÓSTICO Y PROGRESIÓN  
DE LA LEUCEMIA LINFOBLÁSTICA AGUDA DE TIPO T  
(LLA-T) DEL ADULTO

Memoria de Tesis presentada por

**Celia González Gil**

Para optar al grado de Doctora en Biomedicina  
por la Universidad de Barcelona

Investigación realizada en

Instituto Josep Carreras

Trabajo realizado bajo la dirección de la Dra. Eulàlia Genescà y  
co-dirección del Dr. Josep María Ribera

Directora:  
Dra. Eulàlia Genescà Ferrer

Co-Director:  
Dr. Josep Maria Ribera  
Santanusana

Tutora:  
Dra. Neus Agell Jane

Celia González Gil

Barcelona,  
junio 2023

Este trabajo de Tesis Doctoral ha sido posible gracias a la financiación recibida:

Financiación para el salario de Celia González Gil:

- Beca predoctoral de l'AGAUR: ajuts destinats a universitats, centres de recerca i fundacions hospitalàries per a la contractació de personal investigador novell per a l'any 2018 (FI) (Ref. BDNS 362582) (2018-2021)
- Beca de investigación de la Sociedad Española de Hematología (Ref. SEHH 2021) (2022-2024)

Financiación para desarrollar el proyecto:

- Proyecto de Investigación financiado por la Asociación Española contra el Cáncer AECC (GC16173697BIGA)
- Proyecto financiado por los Fondos de Investigación en Salud (FIS) Instituto de Salud Carlos III (PI19/01183)
- Proyecto financiado por los Fondos de Investigación en Salud (FIS) Instituto de Salud Carlos III (PI19/01183)

# ÍNDICE





## ÍNDICE

---

<b>ABREVIATURAS</b> .....	<b>3</b>
<b>INTRODUCCIÓN</b> .....	<b>7</b>
1. HEMATOPOYESIS .....	9
2. LINFOPOYESIS T .....	11
3. LEUCEMIA LINFOBLÁSTICA AGUDA .....	13
3.1. CONCEPTO .....	13
3.2. EPIDEMIOLOGÍA/INCIDENCIA .....	14
3.3. ETIOLOGÍA .....	14
3.4. DIAGNÓSTICO .....	16
3.4.1 Cuadro clínico .....	16
3.4.2 Técnicas de diagnóstico .....	16
3.5. CLASIFICACIÓN DE LA LLA-T .....	20
3.5.1 Clasificación morfológica .....	20
3.5.2 Clasificación inmunofenotípica .....	20
3.5.3 Clasificación de la OMS .....	21
3.5.4 Clasificación de la ICC .....	21
3.5.5 Clasificación genética .....	22
3.5.5.1 Subtipos genéticos según factor de transcripción afectado .....	23
3.5.5.2 Alteraciones genéticas secundarias .....	28
3.6. FACTORES PRONÓSTICOS EN LA LLA-T .....	41
3.6.1 Factores pronósticos del paciente adulto .....	42
3.6.2 Factores pronósticos de la enfermedad .....	43
3.6.3 Factores pronósticos de respuesta al tratamiento .....	44
3.7. TRATAMIENTO .....	46
3.8. EVOLUCIÓN CLONAL Y RECAÍDA .....	48
3.7.1 Evolución clonal .....	49
3.7.2 Recaída .....	52
3.7.2.1 Mutaciones de recaída .....	52
3.7.2.2 Factores pronósticos de la recaída .....	54
3.7.2.3 Nuevos agentes para pacientes recaídos/refractarios .....	54
<b>HIPÓTESIS Y OBJETIVOS</b> .....	<b>61</b>
<b>METODOLOGÍA</b> .....	<b>64</b>
<b>METODOLOGÍA DEL OBJETIVO 1</b> .....	<b>66</b>

1. Pacientes y diseño del estudio .....	66
1.1. Procesamiento de las muestras, extracción del ADN y control de calidad .....	66
1.1.1. Descongelación .....	67
1.1.2. Enriquecimiento de las muestras en contenido de blastos T ( <i>sorting</i> ) .....	67
1.1.3. Extracción del ADN y control de calidad.....	67
1.2. Diseño del panel de secuenciación y validación .....	68
1.2.1. Diseño del panel de secuenciación .....	68
1.2.2. Validación .....	70
1.3. Detección y análisis de las alteraciones tipo SNV/indel.....	72
1.3.1. Preparación de las librerías, enriquecimiento por hibridación con sondas y captura.....	72
1.3.2. Secuenciación .....	72
1.3.3. Procesamiento de los datos secuenciados .....	72
1.3.4. Filtrado de las variantes .....	73
1.3.5. Comparación del perfil mutacional de muestras pareadas sangre periférica y médula ósea.....	75
1.4. Estudio de la variación del número de copias.....	76
1.5. Corrección de la frecuencia alélica y estimación de la jerarquía clonal. ....	78
1.6. Asociaciones entre mutaciones y características biológicas.....	81
<b>METODOLOGÍA DEL OBJETIVO 2.....</b>	<b>82</b>
1. Pacientes y diseño del estudio .....	82
2. Esquemas de tratamiento.....	83
3. Análisis estadísticos .....	83
<b>METODOLOGÍA DEL OBJETIVO 3.....</b>	<b>86</b>
1. Pacientes y diseño del estudio .....	86
2. Detección y análisis de las alteraciones tipo SNV/indel .....	86
3. Análisis de las variaciones en el número de copias (CNV) .....	86
4. Análisis y seguimiento de mutaciones con PCR digital en muestras seriadas .....	88
5. Amplificación de leucemias LLA-T primarias en ratones inmunodeprimidos (PDX) ..	91
6. Análisis estadísticos .....	93
<b>RESULTADOS.....</b>	<b>95</b>
<b>RESULTADOS DEL OBJETIVO 1.....</b>	<b>97</b>
1. Caracterización genómica de los pacientes con LLA-T en el momento del diagnóstico. ....	97
1.1. Alteraciones detectadas por TDS .....	97

1.2. Alteraciones detectadas por SNP- <i>arrays</i> y determinación de los genes diana en las alteraciones tipo CNV .....	99
1.3. Panorama mutacional de la LLA-T .....	110
1.4. Cooperación entre las alteraciones genéticas en el desarrollo de la LLA-T y en la definición de las características biológicas de los pacientes.....	113
1.5. Evaluación de la jerarquía clonal de la LLA-T .....	114
<b>RESULTADOS DEL OBJETIVO 2.....</b>	<b>117</b>
1. Impacto pronóstico de las alteraciones detectadas por TDS y SNP- <i>arrays</i> .....	117
1.1. Impacto en la supervivencia global de las alteraciones tipo SNV/indel .....	117
1.1.1. Impacto en la supervivencia global de las mutaciones que afectan a genes de la vía de señalización de NOTCH1 ( <i>NOTCH1</i> y <i>FBXW7</i> ).....	118
1.1.2. Las mutaciones en <i>DNMT3A</i> , <i>U2AF1</i> , <i>N/KRAS</i> y/o <i>MSH2</i> identifican a pacientes con peor supervivencia global.....	120
1.2. Impacto en la supervivencia global de las alteraciones tipo CNV .....	127
1.2.1. Deleciones asociadas a peor supervivencia global de los pacientes con LLA-T .....	127
1.2.2. Impacto en la supervivencia global de las deleciones en <i>CDKN2A/B</i> .....	129
1.2.3. Impacto en la supervivencia global de las alteraciones en <i>PTEN</i> .....	130
1.3. Impacto de las alteraciones en la incidencia acumulada de recaída de los pacientes con LLA-T .....	130
<b>RESULTADOS DEL OBJETIVO 3.....</b>	<b>132</b>
1. Comparación de las alteraciones genéticas entre el diagnóstico y la recaída.....	132
1.1. Diferencias en el perfil mutacional de las muestras pareadas .....	135
2. Identificación y seguimiento de las mutaciones en <i>NT5C2</i> mediante dPCR.....	141
3. Implicación del tratamiento en la expansión de los clones con mutaciones específicas de recaída .....	144
<b>DISCUSIÓN .....</b>	<b>145</b>
<b>CONCLUSIONES.....</b>	<b>145</b>
<b>BIBLIOGRAFÍA .....</b>	<b>145</b>
<b>ANEXOS.....</b>	<b>145</b>





# ABREVIATURAS



- aa:** aminoácido
- ADN:** ácido dexoxirribonucleico
- AGM:** aorta-gónada-mesonefros
- Ara-C:** arabinósido de citosina
- ARN:** ácido ribonucleico
- AYA:** adolescentes y adultos jóvenes
- BFLS:** *Börjeson–Forssman–Lehmann syndrome*
- BFM:** *Berlin- Frankfurt- Münster*
- CAR:** *chimeric antigen receptor* (receptor de antígeno quimérico)
- CCF:** *cancer cell fraction* (fracción celular del cáncer)
- CD:** *cluster of differentiation* (cluster de diferenciación)
- CDK:** *cyclin-dependent kinases* (ciclina dependiente de quinasa)
- CDP:** *common dendritic progenitor* (progenitor común dendrítico)
- CFM:** citometría de flujo multiparamétrica
- CGH:** *comparative genome hybridization* (hibridación genómica comparada)
- CH:** *clonal hematopoiesis* (hematopoyesis clonal)
- CHIP:** *clonal hematopoiesis of indeterminate potential* (hematopoyesis clonal de potencial indeterminado)
- CIR:** *cumulative incidence of relapse* (incidencia acumulada de recaída)
- CLP:** *common lymphoid progenitors* (progenitores linfoides comunes)
- CMJ:** *junction corticomedullary* (unión cortico-medular)
- CMP:** *common myeloid progenitor* (progenitor común mieloide)
- CNV:** *copy number variation* (variación en el número de copias)
- del:** delección
- DIN:** *DNA integrity number* (Número de integridad del ADN)
- DP:** doble positiva
- dPCR:** PCR digital
- dup:** duplicación
- DX:** diagnóstico
- ECOG:** *Eastern Cooperative Oncology Group* (Grupo cooperativo de oncología del Este)
- EC:** endothelial cells (células endoteliales)

**EGIL:** *European group for the immunological characterization of leukemias* (grupo europeo para la clasificación inmunológica de las leucemias)

**ELP:** *early lymphoid progenitor* (progenitor linfóide temprano)

**EMA:** *European Medicines Agency* (Agencia Europea de Medicamentos)

**ERM:** enfermedad residual medible

**ESMO:** *European Society for Medical Oncology* (sociedad europea de medicina oncológica)

**ETP:** *early T-cell precursors* (células de progenitores tempranos)

**EWALL:** *European Working Group for Adult Acute Lymphoblastic Leukemia*, (grupo de trabajo europeo de LLA adulto)

**FAB:** Francés-Americano-Británico

**FBS:** *fetal bovine serum* (suero bovino fetal)

**FDA:** *Food and Drug Administration* (Administración de Alimentos y Medicamentos)

**FISH:** *fluorescence in situ hybridization* (hibridación *in situ* fluorescente)

**FT:** factor de transcripción

**GMALL:** *German Multicenter Study Group for Adult ALL* (Grupo multicéntrico alemán del estudio de la LLA en adultos)

**GMP:** *granulocyte-macrophage progenitor* (progenitor de granulocitos y macrófagos)

**GOG:** *good outcome genetics* (genética de buen pronóstico)

**GRAALL:** *Grupo de Recherche sur les Acutés Lymphoblastiques de l'Adulte* (Grupo Francés para la Investigación en Leucemia Aguda Linfoblástica)

**GSI:** *gamma-secretase inhibitors* (inhibidores de gamma-secretasa)

**GWAS:** *genome-wide association studies* (estudios de asociación de todo el genoma)

**HR:** *Hazard Ratio*

**HSC:** *hematopoietic stem cell* (célula madre hematopoyética)

**IC95%:** Intervalo de confianza del 95%

**ICN1:** *intracellular domain of Notch* (dominio intracelular de Notch)

**Ig:** Inmunoglobulina

**IGTP:** Instituto Germans Trias i Pujol

**IL-7:** interleucina 7

**IL7R:** *interleukin-7 receptor* (receptor de la interleucina 7)

**indel:** variantes de inserción y deleción

**ISP:** *immature single positive* (células inmaduras simples positivas)

**ITD:** *internal tandem duplications* (duplicaciones en tándem internas)

- IT-HSC:** *intermediate term HSC* (célula madre hematopoyética de plazo intermedio)
- LAIP:** *leukemia-associated immunophenotypes* (inmunofenotipo asociado leucemia)
- LCR:** líquido cefalorraquídeo
- LIC:** *leukemic initiating cells* (células iniciadoras de la leucemia)
- LLA-B:** leucemia linfoblástica aguda de tipo B
- LLA-T:** leucemia linfoblástica aguda de tipo T
- LMA:** leucemia mieloide aguda
- LMC:** leucemia mieloide crónica
- LMPP:** *lymphoid-primed MPP* (MPP de pre-instrucción linfoide)
- LOD:** *limit of detection* (límite de detección)
- LOH:** *loss of heterocigosity* (pérdidas de heterocigosidad )
- LRF:** *leukemia/lymphoma-related factor* (factor de transcripción relacionado con leucemia/linfoma)
- LT-HSC:** *long-term hematopoietic stem cell* (célula madre hematopoyética a largo plazo)
- MDP:** *monocyte-dendritic cell progenitor* (progenitor de monocitos y células dendríticas)
- MEP:** *megakaryocyte/erythroid progenitor* (progenitor megacariocito/eritroide)
- MHC:** *major histocompatibility complex* (complejo mayor de histocompatibilidad)
- MO:** médula ósea
- MPAL:** *mixed phenotype acute leukemia* (leucemias de linaje mixto)
- MPO:** mieloperoxidasa
- MPP:** *multipotent progenitors* (progenitores multipotentes)
- NGS:** *next generation sequencing* (secuenciación de nueva generación)
- NK:** Natura Killer
- NOPHO:** *Nordic Society of Paediatric Haematology and Oncology* (sociedad nórdica de hematología y oncología pediátrica)
- OMS:** Organización Mundial de la Salud
- OR:** *odds ratio*
- OS:** *overall survival* (supervivencia global)
- pb:** pares de base
- PBS:** *phosphate buffered saline* (Tampón fosfato salino)
- PCR:** *polymerase chain reaction* (reacción en cadena de la polimerasa)
- PDX:** *patient-derived xenograft* (paciente derivado de xenoinjerto)

**PETHEMA:** Programa Español de Tratamientos en Hematología

**Ph:** *Philadelphia chromosome* (cromosoma Filadelfia)

**PRC2:** *polycomb repressive complex 2* (complejo represivo Polycomb 2)

**Prob:** probabilidad

**qPCR:** *real time- quantitative polymerase chain reaction* (reacción en cadena de la polimerasa en tiempo real)

**RBC:** *red blood cells* (glóbulos rojos)

**RC:** remisión completa

**RE:** recaída

**RT-PCR:** *reverse transcription polymerase chain reaction* (reacción en cadena de la polimerasa de transcripción inversa)

**SMD:** síndrome mielodisplásico

**SNC:** sistema nervioso central

**SNP:** *single nucleotide polymorphism* (polimorfismos de único nucleótido)

**SNV:** *single nucleotide variant* (variante de un único nucleótido)

**SP:** sangre periférica

**ST-HSC:** *short-term hematopoietic stem cell* (célula madre hematopoyética a largo plazo)

**TCR:** *T-cell receptor* (receptor de las células T)

**TDS:** *target deep sequencing* (secuenciación dirigida)

**TIT:** triple intratecal

**TKI:** *tyrosin kinase inhibitor* (inhibidor tirosina quinasa)

**TPH:** Trasplante de progenitores hematopoyéticos

**UKALL:** *United Kingdom Acute Lymphoblastic Leukemia* (Leucemia Linfoblástica Aguda del Reino Unido)

**VAF:** *variant allele frequency* (frecuencia alélica de la variante)

**WBC:** *white blood cells* (recuento de glóbulos blancos)

**WES:** whole exome sequencing (secuenciación del exoma completo)

**WGS:** *whole genome sequencing* (secuenciación del genoma completo)

**WOG:** *Worse outcome genetics* (genética de mal pronóstico)

**WT:** *wild type* (no alterado)



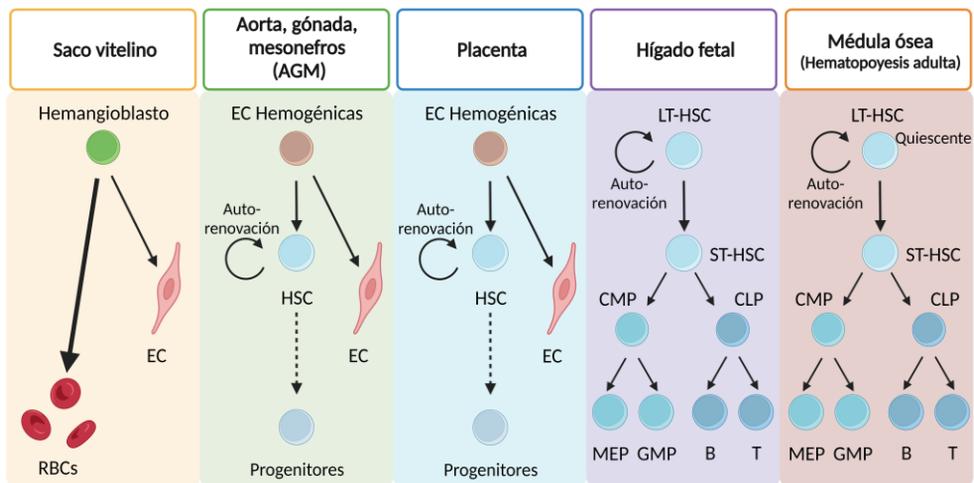
# INTRODUCCIÓN



## 1. HEMATOPOYESIS

La hematopoyesis es el proceso fisiológico por el cual las células que componen la sangre se forman y renuevan continuamente a lo largo de la vida de un organismo. El sistema hematopoyético contiene más de 10 tipos diferentes de poblaciones celulares altamente especializadas, jerarquizadas y con funciones únicas. Es uno de los sistemas más regenerativos y plásticos del cuerpo humano y millones de sus células se reemplazan por nuevas cada segundo durante la vida, lo que requiere que esté altamente regulado<sup>1</sup>.

El desarrollo hematopoyético embrionario ocurre en varias etapas, en las cuales se van formando de manera secuencial los diferentes tipos celulares (Fig. 1). Existe una primera etapa, denominada etapa primitiva, donde se generan los eritrocitos y macrófagos a partir del hemangioblasto, precursor común con las células endoteliales. Las células que se producen en esta etapa no son multipotentes y no tienen capacidad de autorrenovación. Se trata de una etapa transitoria que ocurre durante el desarrollo temprano del embrión, en la región extraembrionaria del saco vitelino, con el fin de proveer de oxígeno al embrión<sup>2,3</sup>.

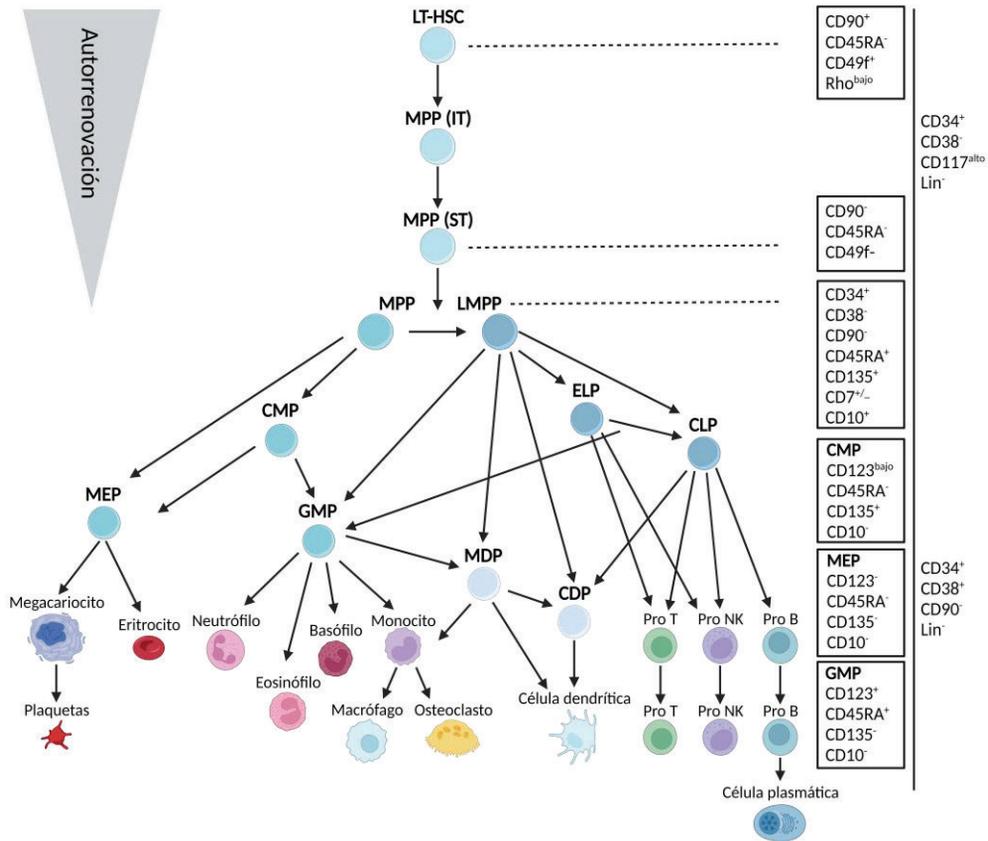


**Figura 1. Desarrollo hematopoyético embrionario.** ECs, *endothelial cells* (células endoteliales); RBCs, *red blood cells* (glóbulos rojos); LT-HSC, *long-term hematopoietic stem cell* (célula madre hematopoyética a largo plazo); ST-HSC, *short-term hematopoietic stem cell* (célula madre hematopoyética de corto plazo); CMP, *common myeloid progenitor* (progenitor común mieloide); CLP, *common lymphoid progenitor* (progenitor común linfocitario); MEP, *megakaryocyte/erythroid progenitor* (progenitor megacariocito/eritroide); GMP, *granulocyte/macrophage progenitor* (progenitor de granulocitos y macrófagos). Adaptada de Orkin y Zon, 2008<sup>4</sup>.

La hematopoyesis definitiva, ocurre en etapas más tardías del desarrollo e implica a las células madre hematopoyéticas (en inglés *Hematopoietic Stem Cell*, HSC) que tienen capacidad de autorrenovación y de generar todas las células sanguíneas de un organismo adulto. Las HSC definitivas se producen en la región aorta-gónada-mesonefros (AGM) a partir de las denominadas células endoteliales hemogénicas<sup>5,6</sup>. De ahí, migran hacia la placenta y al hígado fetal, que son los mayores sitios de expansión. Finalmente, colonizan el

timo y próximo al nacimiento se establecen en la médula ósea, lugar donde se produce la hematopoyesis en los adultos<sup>1</sup>.

En la hematopoyesis es básico el mantenimiento del número de cada tipo celular teniendo que existir un equilibrio entre la autorrenovación y la diferenciación. Los modelos de jerarquía hematopoyética actuales afirman que la generación de los diferentes linajes hematopoyéticos se basa en la pérdida sucesiva de potenciales de linaje individuales, haciendo cada linaje inferior más diferenciado y comprometido (Fig. 2).



**Figura 2. Jerarquía celular de la hematopoyesis humana.** LT-HSC, *long-term hematopoietic stem cell* (célula madre hematopoyética a largo plazo); MPP, *multipotent progenitor* (progenitor multipotente); LMPP, *lymphoid-primed MPP* (MPP de pre-instrucción linfoide); ELP, *early lymphoid progenitor* (progenitor linfoide temprano); CLP, *common lymphoid progenitor* (progenitor común linfoide); CMP, *common myeloid progenitor* (progenitor común mieloide); MEP, *megakaryocyte-erythrocyte progenitor* (progenitor de megacariocitos y eritrocitos); GMP, *granulocyte-macrophage progenitor* (progenitor de granulocitos y macrófagos); MDP, *monocyte-dendritic cell progenitor* (progenitor de monocitos y células dendríticas); CDP, *common dendritic progenitor* (progenitor común dendrítico). Adaptada de Rieger y Schroeder, 2012<sup>1</sup>.

Este proceso está altamente regulado por factores intrínsecos (factores de transcripción, reguladores epigenéticos, reguladores del ciclo celular, microARN, etc) y extrínsecos

(derivados del microambiente del nicho y citoquinas, principalmente) que favorecen la expresión de genes específicos de un linaje determinado e inhiben la expresión de los específicos de otros linajes<sup>1</sup>.

De este modo, según estos modelos, las HSC se encuentran en la cúspide de la jerarquía hematopoyética en un estado quiescente y cuando se activan se inicia el proceso de diferenciación celular. Las HSC son las únicas células con potencial de autorrenovación y diferenciación a largo plazo (en inglés *long term HSCs progenitors*, LT-HSC). A partir de estas se origina la población de progenitores con capacidad de repoblación intermedia (en inglés *intermediate term HSC*, IT-HSC) y a continuación, los progenitores con capacidad de repoblación corta y finita (en inglés *short term HSC*, ST-HSC). De estos últimos, derivan los progenitores multipotentes (en inglés, *multipotent progenitors*, MPP), que conservan la capacidad de dar lugar a diferentes linajes hematopoyéticos y se diferencian en progenitores comunes mieloides (en inglés, *common myeloid progenitors*, CMP) y en progenitores comunes linfoides (en inglés, *common lymphoid progenitors*, CLP). De los CMP derivan las células del linaje mieloides: monocitos, granulocitos, eritrocitos y megacariocitos. De los CLP derivan las células del linaje linfoides: linfocitos T, linfocitos B y células *Natural Killer* (NK)<sup>1</sup>.

No obstante, la diferenciación hematopoyética es mucho más compleja y las mismas células de un linaje pueden originarse a partir de distintos progenitores y vías de diferenciación. Un ejemplo de esto son los precursores de células T tempranas (en inglés, *early T-cell precursors*, ETP) que mantienen cierto potencial de NK, células dendríticas y mieloides, y se ha descrito que pueden desarrollarse por una vía independiente a la de los CLP<sup>7</sup>.

## 2. LINFOPOYESIS T

En esta tesis se estudia en concreto la leucemia linfoblástica aguda de tipo T (LLA-T) y para entender el origen y desarrollo de ésta, es básico comprender como se produce el proceso de maduración y diferenciación de los linfocitos T. Además, las características de cada una de estas células ofrecen gran información acerca de las características que tienen sus correspondientes leucémicas y los diferentes subtipos de LLA-T.

La diferenciación de los linfocitos T es distinta a la del resto de células hematopoyéticas, pues en lugar de desarrollarse en la médula ósea, el proceso se completa en el timo, ya que la diferenciación de las células T dependen de señales extrínsecas liberadas por este órgano<sup>8,9</sup>. En concreto, el timo libera citoquinas (CCL21 y CCL25) que movilizan desde la médula a los progenitores MPP y CLP que expresan receptores para estas (CCR7 y CCR9)<sup>10,11</sup>. Estos progenitores entran al timo por la unión cortico-medular (en inglés *junction corticomedullary*, CMJ) y van madurando a través de este órgano mediante la expresión de programas transcripcionales específicos.

En este proceso es fundamental el papel del receptor NOTCH1 que se activa por la unión de ligandos específicos que expresan las células endoteliales corticales del timo, y promueve la diferenciación de los precursores T, haciendo que éstos pierdan de forma gradual e irreversible su potencial de convertirse en otros linajes hematopoyéticos<sup>12-14</sup>. De este modo la activación de NOTCH1 desencadena: i) la conversión de los progenitores LMPP en células

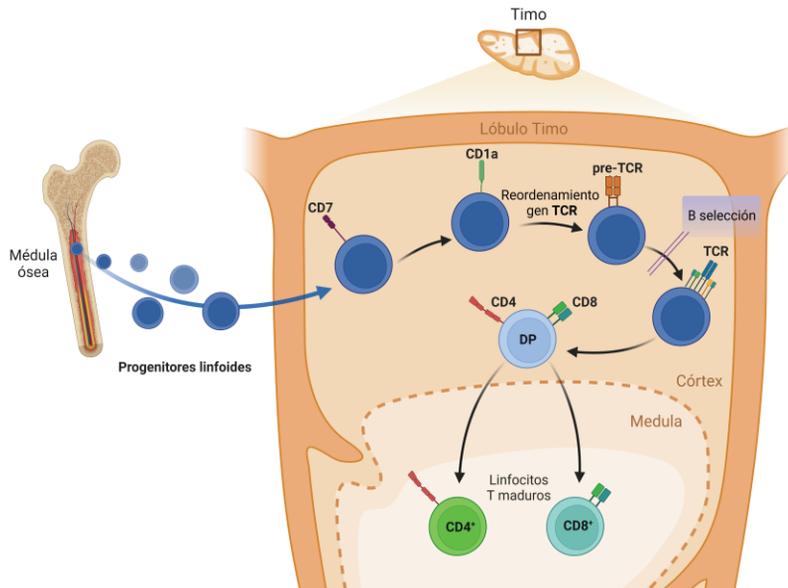
progenitoras tempranas T, inicialmente mediante la activación de TCF-1 y posteriormente de GATA3 y BCL11B<sup>15,16</sup>; ii) la inhibición de la expresión de EBF1 y E2A responsables de la diferenciación hacia células de linaje B<sup>17,18</sup>, iii) la disminución del potencial mielóide mediante la inhibición de los componentes responsables de este proceso (PU.1 y C/EBP $\alpha$ )<sup>19</sup>, iv) el bloqueo de la expresión del componente que regula el desarrollo de las células NK (ID2)<sup>20</sup>. En cambio, la expresión de NOTCH1 está inhibida en la médula ósea por parte del factor de transcripción relacionado con leucemia/linfoma (en inglés, *leukemia/lymphoma-related factor*, LRF), lo que evita el desarrollo de los linfocitos T en la médula ósea<sup>21</sup>.

Los precursores T a través de su paso por el timo van cambiando los marcadores que expresan en su superficie. En los humanos se conoce la transición de estos marcadores, pero se desconoce el punto exacto del timo donde ocurre cada una. Tampoco se conoce la identidad exacta, a nivel de marcadores de superficie, de las células ETP, que llegan al timo desde la médula ósea. No obstante, los progenitores más tempranos detectados en el timo humano son células CD34<sup>+</sup>, CD7<sup>+</sup>, CD1a<sup>22</sup>. Por tanto, CD7 es el marcador de células T expresado de forma más temprana en los timocitos humanos. Posteriormente, se adquiere la expresión de CD5, seguido de la expresión de CD1a<sup>23</sup>, momento en el cual las células T ya no son capaces de formar células de otros linajes<sup>22-25</sup>. En este estadio también se produce el reordenamiento de la cadena  $\beta$  del receptor de las células T (en inglés *T-cell receptor*, TCR), mediado por los genes RAG-1 y RAG-2<sup>26,27</sup>. El siguiente estadio de maduración se caracteriza por la expresión del marcador CD4, en ausencia del marcador CD8, lo que se conoce como células inmaduras simple positivas (en inglés *immature single-positive*, SP) (CD4ISP). Debido a que las cadenas  $\alpha$  del TCR aún no se han reordenado, las cadenas  $\beta$  se unen covalentemente a las cadenas del pre-TCR $\alpha$ , formándose lo que se conoce como pre-TCR. A continuación, estos linfocitos, aún no maduros, sufren el proceso de la  $\beta$ -selección, donde aquellas células que han reordenado correctamente se seleccionan para continuar el proceso de diferenciación<sup>28</sup>.

Las señales pre-TCR promueven la supervivencia, proliferación y diferenciación de las células CD4ISP a células doble positivas (DP) co-expresándose los receptores CD4 y CD8 en su superficie, junto con el complejo CD3. Finalmente, se produce el reordenamiento de las cadenas  $\alpha$  del TCR<sup>26,29,30</sup>, que se unen con las cadenas  $\beta$  generándose el TCR $\alpha\beta$  expresado en las células DP. En este momento se produce la selección positiva y negativa de las células DP que reconocen antígenos propios presentados por el complejo mayor de histocompatibilidad (en inglés *major histocompatibility complex*, MHC). Por último, las células DP seleccionadas se diferencian hacia células simple positivas (SP) que dejan de expresar CD1a<sup>31</sup> y pasan a ser células CD4<sup>+</sup> o células CD8<sup>+</sup> localizadas en la médula del timo. En la figura 3 se representa este proceso de diferenciación.

Otro componente clave en el proceso de diferenciación de las células T es la interleucina 7 (IL-7), producida por las células corticales y medulares del timo. Su función varía en las diferentes etapas de la diferenciación y está regulada por la expresión de su receptor, IL-7R. La vía de señalización activada por IL-7, tras la unión al IL-7R, favorece la supervivencia y expansión de los progenitores linfoides<sup>32,33</sup>, la proliferación de los timocitos antes de que reordenen las cadenas  $\beta$  del TCR y regula la duración de esta etapa<sup>34,35</sup>. Una vez constituido

el pre-TCR, la expresión de *IL7R* (concretamente de *IL7R $\alpha$* ) disminuye hasta desaparecer en los timocitos DP<sup>36-38</sup>. Tras la selección positiva se reestablece<sup>38</sup>, asegurando la proliferación de las células T maduras periféricas derivadas de los timocitos SP (CD4+ /CD8+)<sup>39</sup>.



**Figura 3. Proceso de diferenciación de las células T.**

Este conjunto de información muestra que la diferenciación de los linfocitos T es un proceso altamente regulado. La alteración del equilibrio homeostático que puede ocurrir durante este proceso es una característica de las leucemias, en este caso concreto de tipo T. Por ello, la presencia de mutaciones en los principales factores de transcripción hematopoyéticos y en los componentes de las vías de señalización que regulan estos procesos de diferenciación son una de las principales causas del desarrollo de la leucemia. Esto se ejemplifica en que las mutaciones somáticas con mayor frecuencia en las leucemias T son las que afectan al gen que codifica para NOTCH1<sup>40</sup>.

### 3. LEUCEMIA LINFOBLÁSTICA AGUDA

#### 3.1. CONCEPTO

La LLA es una neoplasia hematológica clonal que se caracteriza por un bloqueo en la diferenciación de los linfocitos inmaduros o linfoblastos y una proliferación anormal de estos precursores que invaden la médula ósea, la sangre y otros órganos con capacidad hematopoyética, desplazando la hematopoyesis sana. Todo ello conduce a una insuficiencia medular, manifestación clínica más destacada de las leucemias agudas.

Existen 2 subtipos principales de LLA en función del progenitor linfoide afectado: LLA de precursores B (LLA-B) que comprende el 75% de los casos en adultos, y LLA de precursores T (LLA-T) que comprende el 25% de los casos en adultos<sup>41</sup>.

### 3.2. EPIDEMIOLOGÍA/INCIDENCIA

La incidencia de la LLA es de 1,8 casos/100.000 habitantes por año. La mediana de edad al diagnóstico es de 17 años, aunque la incidencia es variable dependiendo del rango de edad. Es el cáncer más común en la edad pediátrica, donde representa una tercera parte de los cánceres infantiles, mientras que su incidencia en la población adulta es baja. Entre el 50-60% de los pacientes con LLA se diagnostican antes de los 20 años, con un pico máximo de incidencia entre 1-4 años. Predomina en pacientes de género masculino y presenta una distribución desigual entre las diferentes etnias<sup>42</sup> (Fig. 4).

En el caso concreto de la LLA-T representa el 10-15% de los casos pediátricos y 20-25% de los casos adultos de LLA en Europa, Estados Unidos y Japón y es el doble de prevalente en hombres que en mujeres<sup>43,44</sup>.



**Figura 4. Distribución de la LLA en la población.** A) Distribución por edad. B) Distribución por sexo<sup>45</sup>

### 3.3. ETIOLOGÍA

La LLA, como un cáncer en general, parece originarse por la combinación de la exposición a agentes exógenos (p.ej. infecciones) y endógenos (p.ej. inflamación o estrés oxidativo), la presencia de variantes hereditarias (susceptibilidad genética) y eventos casuísticos<sup>46</sup>. No obstante, para la mayoría de las LLA se desconocen las causas exactas de su origen. Además, existen diversos subtipos y puede que no todos tengan los mismos mecanismos etiológicos<sup>46</sup>. De hecho, el origen de una leucemia pediátrica puede ser diferente al que se da en los adultos, como se ejemplifica a continuación.

Hay estudios que han mostrado que la exposición a agentes físicos y químicos (p.ej. radiaciones ionizantes y no ionizantes, pesticidas, etc.) tienen potencial leucemogénico<sup>47-49</sup>. No obstante, una de las primeras causas exógenas propuestas para explicar el origen de la LLA pediátrica fue la infección por virus o bacterias, y actualmente continúa siendo una de las principales teorías. Hay dos hipótesis que explican esta asociación, y ambas coinciden en que la LLA podría ser consecuencia de una respuesta inmune anormal a una infección vírica o bacteriana. Se trata de la hipótesis de Kinlen y la de Greaves<sup>50,51</sup>. La hipótesis de Kinlen indica que el contacto de los virus con la población pediátrica, poco expuesta a estas

infecciones, desencadenaría una respuesta aberrante del sistema inmune y la aparición de leucemias infantiles<sup>50</sup>. Por su parte, Greaves propone que el desarrollo de la LLA se debe, como mínimo, a la presencia de dos eventos mutacionales. El primero (*first hit*) tendría lugar en el útero o en una etapa muy temprana posnatal, como por ejemplo ocurre con las traslocaciones de *KMT2A*, la t(12;21)(p13;q22) (fusión *ETV6::RUNX1*) y algunas aneuploidías, todas descritas por tener un origen intrauterino<sup>52,53</sup>. Posteriormente, la exposición a una infección desencadenaría una excesiva proliferación linfocitaria que facilitaría la aparición de segundas mutaciones (*second hit*)<sup>54</sup>. Hay varios estudios que apoyan la hipótesis de Greaves, como los estudios realizados en gemelos univitelinos, donde ambos nacen con este primer evento leucemogénico, pero solo el gemelo que padece leucemia es portador de otras alteraciones genéticas posnatales, lo que indicaría la necesidad de ambos eventos mutacionales (*doble hit*) para el desarrollo de la leucemia<sup>55-57</sup>.

Por otra parte, se sabe que una pequeña proporción de casos pediátricos de LLA están asociados a síndromes genéticos, como el síndrome de Down, la anemia de Fanconi, el síndrome Bloom, la ataxia-telangiectasia y el síndrome de rotura de Nijmegen<sup>58-61</sup>, lo que sugiere una relación de predisposición. Además, cada vez hay más evidencias de la presencia de variantes germinales de predisposición a la LLA<sup>62</sup>. En el caso de la LLA-T, mediante estudios de asociación de todo el genoma (en inglés *genome-wide association studies*, GWAS) en pacientes pediátricos, se han identificado polimorfismos que afectan a un único nucleótido (en inglés *single nucleotide polymorphism*, SNP) en los genes *CDKN2A*, *USP7*, *IKZF1* y *RUNX1*, que aumentan el riesgo de padecer esta enfermedad<sup>63-66</sup>. Todos estos genes están implicados en procesos de regulación del desarrollo, proliferación y diferenciación de las células hematopoyéticas. En ellos, también se identifican mutaciones somáticas en la LLA, lo que sugiere que las variantes genéticas heredadas hacen más vulnerables a los progenitores hematopoyéticos a sufrir eventos de transformación, ya sea en el útero o en estadios posnatales tempranos. Aunque la predisposición germinal siempre se ha asociado a casos pediátricos, cada vez hay más evidencias de que el cáncer en pacientes adultos también podría tener un componente germinal<sup>67</sup>.

Sin embargo, la etiología de la LLA en general y LLA-T en particular, es aún más desconocida en el caso de los pacientes adultos. Es un hecho que los humanos van adquiriendo mutaciones no patogénicas durante su vida. En concreto, se estima que una célula hematopoyética normal puede adquirir aproximadamente 1 mutación por década de vida<sup>68</sup>. Algunas de estas mutaciones pueden conferir una ventaja adaptativa dando lugar a una proliferación clonal, fenómeno denominado hematopoyesis clonal (en inglés *clonal hematopoiesis*, CH). Estudios de secuenciación han identificado en individuos sanos la presencia de mutaciones en genes implicados en la patogénesis de neoplasias hematológicas, especialmente leucemia mieloide aguda (LMA), a unas frecuencias alélicas (en inglés *variant frequency allelic*, VAF) muy bajas, de alrededor del 2%. Estas mutaciones se consideran como un evento pre-leucémico y a este proceso se le denomina hematopoyesis clonal de potencial indeterminado (en inglés *clonal hematopoiesis of indeterminate potential*, CHIP)<sup>69</sup>. En el caso de la LLA-T, se han identificado mutaciones en el gen *DNMT3A*, uno de los genes más frecuentemente alterado en los casos de CHIP, en células de origen no leucémico en pacientes adultos con un inmunofenotipo inmaduro<sup>70</sup>, lo

que sugiere que estas LLA-T más inmaduras podrían originarse a partir de eventos de tipo CHIP.

No obstante, aunque este conjunto de evidencias muestran la importancia de los eventos genéticos en el origen de la LLA, su etiología también está gobernada por la desregulación de la expresión génica debida a procesos epigenéticos como la metilación, cambios en el patrón de las histonas, ARN no codificantes, entre otros; además de las interacciones de las células pre-leucémicas con las células del microambiente<sup>71,72</sup>.

### 3.4. DIAGNÓSTICO

#### 3.4.1 Cuadro clínico

La mayoría de las manifestaciones clínicas que presenta un paciente con LLA son un reflejo de la acumulación de blastos leucémicos en la médula ósea, sangre periférica y sitios extramedulares. La presentación clínica puede ser inespecífica, con una combinación de síntomas constitucionales y signos derivados de una insuficiencia medular (anemia, trombocitopenia, leucopenia). Los síntomas comunes incluyen fiebre, pérdida de peso, hemorragias, fatiga, disnea e infecciones. Es común la afectación de sitios extramedulares y puede cursar con adenopatías, esplenomegalia o hepatomegalia<sup>41</sup>. La afectación del sistema nervioso central (SNC) en el momento del diagnóstico ocurre en el 5-8 % de los pacientes<sup>73</sup>, aunque en el caso de la LLA-T este porcentaje es mayor (10%)<sup>74</sup>. De manera característica, la LLA-T aparece en forma de masas en la zona del mediastino en la mitad de los casos<sup>41</sup>.

#### 3.4.2 Técnicas de diagnóstico

El diagnóstico de una LLA es un proceso integrado que incluye el análisis morfológico, inmunofenotípico, molecular y genético/citogenético, teniendo en cuenta los criterios y las recomendaciones del actual sistema de clasificación establecido por la organización mundial de la salud (OMS) en 2022<sup>75</sup> y el protocolo de tratamiento a aplicar. Para hacer estos análisis se realiza un aspirado de médula ósea. En la tabla 1 se resumen las aproximaciones diagnósticas de la LLA según las recomendaciones del grupo de trabajo europeo de LLA adulto (en inglés *European Working Group for Adult Acute Lymphoblastic Leukemia, EWALL*)<sup>76</sup>.

En primer lugar, se realiza una evaluación morfológica, mediante microscopía óptica, para determinar el grado de infiltración en la médula ósea y la sangre periférica. Por definición, la médula ósea debe tener al menos un 20% de infiltración de linfoblastos para definirse como una LLA<sup>77</sup>. Además, es obligatorio el examen del líquido cefalorraquídeo (LCR), para descartar infiltración del SNC.

El análisis del inmunofenotipo se realiza mediante citometría de flujo multiparamétrica (CFM) que permite confirmar el origen linfoide de la población blástica, diferenciar entre un fenotipo de linaje B o T y dentro de cada linaje definir el subtipo de LLA. Además, con la CFM también se estudian determinados marcadores que sirven para monitorizar y hacer un seguimiento de la enfermedad residual medible (ERM).

**Tabla 1. Principales técnicas para el diagnóstico de la LLA del adulto propuestas por las guías ESMO**

Proceso diagnóstico	Resultado/subtipo LLA	Recomendación
<b>Morfología</b>		
Médula ósea y sangre periférica	Blastos linfoides/ indiferenciados ( $\geq 20\%$ infiltración medular)	Obligatorio
Líquido cefalorraquídeo	Infiltración del SNC	Obligatorio
<b>Citometría de flujo</b>		
MPO (diagnóstico diferencial respecto a LMA)	MPO negativa; marcadores B/T $>20\%$ (CD3/CD79a $>10\%$ )	Obligatorio
Marcadores de línea B: CD19, CD79a, cCD22 (al menos 2); otros: TdT, CD10, CD20, CD24, clgM, slg (kappa o lambda)	LLA-B: Pro-B/B-I (CD19/ CD79a/ cCD22) Común/ B-II (CD10+/clgM-) Pre-B/ B-III (clgM+/slg-) B Madura/ B-IV (slg+)	Obligatorio
Marcadores de línea T: cCD3; otros: TdT, CD1a, CD2, CD5, CD7, CD4, CD8, TCR $\alpha/\beta$ o $\gamma/\delta$	LLA-T: Pro-T/ T-I (cCd3/CD7+) Pre-T/ T-II (CD2/CD5) T cortical/ T-III (CD1a+) T madura/ T-IV (CD3+/CD1a-)	Obligatorio
Marcadores de HSC/mieloides: CD34, CD13, CD33, CD117	Variable, utilizado para detectar antígenos aberrantes para el seguimiento de ERM, orientar subsecuentes pruebas genéticas	Recomendación
<b>Citogenética/ Biología Molecular</b>		
Cariotipo/ FISH/ RT-PCR	LLA con características clínico-biológicas adversas: <i>Ph+</i> (rápida detección, terapia con TKI) t(4;11) (q21;q23) ( <i>KMT2A::AF4</i> ) t(1;19) (q23;p13) ( <i>TCF3::PBX1</i> ) Otros cariotipos de alto riesgo	Obligatorio
CGH/ SNP/ GEP/ NGS	LLA con características clínico-biológicas adversas: LLA <i>Ph-like</i> ETP-ALL <i>NOTCH1/FBXW7</i> no mutado & <i>RAS/PTEN</i> mutado (para LLA-T) <i>IKZF1/CRFL2/MLL/TP53/CREBBP/RAS</i>	Recomendable para nuevos ensayos clínicos
<b>Estudio de ERM</b>		
Marcadores de ERM: LAIP (inmunofenotipo) /qPCR	Clasificación del riesgo según ERM	Obligatorio

Adaptada de Hoelzer *et al.* 2016 (Grupo ESMO: *European Society for Medical Oncology*)<sup>76</sup>. LLA: leucemia aguda linfoblástica; SNC: sistema nerviosos central; MPO: mieloperoxidasa; LMA: leucemia mieloide aguda; c: citoplasmático; Ig: inmunoglobulina; s: superficie; FISH: *fluorescence in situ hybridation* (hibridación de fluorescencia *in situ*); RT-PCR: *reverse transcriptase polymerase chain reaction* (reacción en cadena de la polimerasa de transcripción inversa); Ph: *philadelphia* TKI: *tyrosine kinase inhibitor* (inhibidor tirosina quinasa); CGH: *comparative genomic hybridation* (hibridación comparativa genómica); SNP: *single nucleotide polymorphism* (polimorfismo de nucleótido único); GEP: *gene expression profiling* (perfil de expresión génica); NGS: *next generation sequencing* (secuenciación de siguiente generación); ERM: enfermedad residual medible; LAIP: inmunofenotipo asociado a leucemia; qPCR: *quantitative polymerase chain reaction* (reacción en cadena de la polimerasa cuantitativa).

El análisis citogenético mediante cariotipo es también obligatorio en el manejo clínico de la LLA. La hibridación *in situ* fluorescente (en inglés *fluorescence in situ hybridization*, FISH) puede ser complementaria en muchas ocasiones. El cariotipo permite detectar alteraciones numéricas (aneuploidías) y estructurales de hasta 5 Mb. La FISH permite detectar alteraciones estructurales de hasta 150-200 kb. El uso de las técnicas citogenéticas permite estratificar a los pacientes según la clasificación establecida por la OMS<sup>75</sup> en el caso de la LLA-B. No obstante, en la LLA-T, dado que la técnica en sí es poco resolutive, las pruebas citogenéticas no ofrecen una clasificación como tal, aunque sí que en algunos casos aportan información sobre estas alteraciones.

Por último, el estudio molecular mediante la reacción en cadena de la polimerasa de transcripción inversa (en inglés *reverse transcription polymerase chain reaction*, RT-PCR) permite, por una parte, confirmar la presencia de reordenamientos específicos (sobre todo genes de fusión), y por otra, detectar los reordenamientos específicos de los genes de las inmunoglobulinas (Ig) o del TCR, que además de confirmar la clonalidad, también se pueden utilizar para el estudio de la ERM.

Actualmente, en la mayoría de los centros asistenciales se ha implementado el uso de nuevas técnicas moleculares de alto rendimiento, dado que las alteraciones detectadas por estas técnicas tienen un papel fundamental en la biología, pronóstico y respuesta al tratamiento de la LLA, como también guían en el uso de terapias dirigidas. Entre estas destacan las matrices para la detección de variaciones en el número de copias (en inglés, *copy number variation*, CNV) y las técnicas de secuenciación masiva o de nueva generación (en inglés, *next generation sequencing*, NGS) para la detección de mutaciones.

- **Matrices para el estudio de alteraciones en el número de copias**

Una matriz genómica (en inglés, *array*) es una colección de fragmentos de ADN unidos de forma ordenada a una superficie sólida. Su fundamento es medir los niveles de hibridación entre estos fragmentos y el ADN de la muestra de estudio, que está marcado con un fluorocromo. A partir de esto se identifican alteraciones de tipo CNV. La resolución es determinada por la distancia genómica de las sondas y su tamaño y es mayor que la del cariotipo y FISH, detectando alteraciones de hasta 20 kb de tamaño. Además, tiene la ventaja respecto a las técnicas citogenéticas tradicionales que no se requieren células en división, sino que el material de partida es ADN genómico. En cuanto a la sensibilidad, permiten detectar alteraciones presentes en aproximadamente un 20% de las células tumorales<sup>78</sup>.

Existen principalmente dos aproximaciones: *Arrays* de hibridación genómica comparada (en inglés *Comparative Genome Hybridization*, CGH-*arrays*) y SNP-*arrays*. En el caso de los CGH-*arrays* la técnica se basa en una hibridación competitiva entre el ADN de estudio y un ADN control, marcados con fluorocromos distintos. Los desequilibrios debidos a las diferencias en el número de copias dan lugar a diferencias en el espectro fluorescente, que se extrapolan en alteraciones. En el caso de los SNP-*arrays* solo se hibrida en el *array* una muestra, también marcada con un fluorocromo, y es la intensidad de cada fragmento lo que permite identificar las alteraciones. Este tipo de *arrays* contienen dos tipos de sondas, unas

para el estudio de CNV y otras que permiten estudiar determinados SNP, lo que permite detectar pérdidas de heterocigosidad (en inglés *loss of heterocigosity*, LOH)<sup>78</sup>.

En el caso específico de la LLA, es un hecho que mediante el análisis por *arrays* se detectan alteraciones relevantes no detectadas por cariotipo o FISH<sup>79,80</sup>.

- **Técnicas de secuenciación masiva**

La secuenciación se basa en la determinación del orden de los 4 nucleótidos que forman el ADN (A, T, C o G) y permite detectar cambios en una única base o variantes de un único nucleótido (en inglés *single nucleotide variants*, SNV) e inserciones y deleciones (indel) de hasta 30-50 pb, aproximadamente. La secuenciación de Sanger ha sido el método más utilizado durante más de dos décadas. A finales de la década 1990 empiezan a desarrollarse las técnicas de secuenciación masiva o NGS, que permiten la secuenciación de millones de reacciones y de múltiples muestras a la vez, a una gran sensibilidad y rendimiento. Existen diferentes aproximaciones que difieren principalmente en la región del genoma que se va a secuenciar<sup>81</sup>:

- Secuenciación del genoma completo (en inglés *Whole Genome Sequencing*, WGS): Se trata de la secuenciación del genoma entero, tanto de las regiones codificantes como de las no codificantes.
- Secuenciación del exoma completo (en inglés, *Whole Exome Sequencing*, WES): Solo se secuencia la parte codificante del genoma (exoma) que constituye el 1-2% del genoma, pero contiene el 85% de las mutaciones descritas.
- Secuenciación dirigida (en inglés *Target Deep Sequencing*, TDS): Es la secuenciación de regiones concretas del genoma. Dependiendo de las regiones incluidas en el panel y de su tamaño, el tamaño total del panel variará.

Para el uso en diagnóstico la mejor aproximación es un panel dirigido que contenga todos los genes con valor diagnóstico, pronóstico y predictivo. Es la opción más sencilla, con menor coste por muestra y con una interpretación de los resultados más rápida y fácil. Aunque los análisis de NGS en su inicio se centraron en el desarrollo de estudios exploratorios y estaban limitados prácticamente al área de la investigación, actualmente el uso de la NGS en la práctica clínica ya es una realidad.

En el diagnóstico de la LLA en España, el último protocolo de tratamiento del grupo PETHEMA (LAL19) [NCT04179929] incluye el análisis de CNV y NGS para evaluar la presencia de alteraciones específicas que guían en el tratamiento a seguir. En el caso concreto de la LLA-T, es necesario estudiar las deleciones del gen *PTEN* por *SNP-arrays* y la presencia de mutaciones en *NOTCH1*, *FBXW7*, *PTEN*, *KRAS* y *NRAS* por NGS.

Por tanto, actualmente las alteraciones genéticas presentes en la leucemia tienen implicación pronóstica y cada vez más forman parte de los criterios de estratificación de los pacientes de los diferentes protocolos de tratamiento. Así pues, tener una buena caracterización de la leucemia en el diagnóstico es muy importante. En el apartado 3.5.5 se encuentran detalladas las principales alteraciones genéticas y moleculares de la LLA-T.

### 3.5. CLASIFICACIÓN DE LA LLA-T

La clasificación actual de la LLA-T, como cualquier otra neoplasia hematológica, integra información de la morfología, inmunofenotipo y alteraciones genéticas y moleculares. No obstante, durante la historia se han propuesto varios métodos para clasificar las leucemias, que han ido evolucionando de manera paralela a las técnicas de diagnóstico y las alteraciones detectadas.

#### 3.5.1 Clasificación morfológica

La primera clasificación de las leucemias la realizó el grupo Francés-Americano-Británico (en inglés *French-American-British*, FAB) y se basa en la evaluación morfológica de las células. Esta clasificación tiene en cuenta el tamaño celular, la relación núcleo/citoplasma, el estado de la cromatina, la forma del núcleo, la presencia de núcleo visible, la basofilia citoplasmática y la presencia de vacuolas<sup>82</sup>. Permite diferenciar una LLA de un linfoma de Burkitt. No obstante, no contempla una distinción entre LLA-B y LLA-T.

#### 3.5.2 Clasificación inmunofenotípica

La clasificación basada exclusivamente en el estudio de la expresión de los marcadores inmunológicos de superficie fue propuesta en 1995 por el grupo europeo para la clasificación inmunológica de las leucemias (en inglés *European Group for the Immunological Characterization of Leukemias*, EGIL)<sup>83</sup>. En el caso de la LLA-T queda dividida en 4 grupos: Pro-T, Pre-T, tímica cortical (T-III) y tímica madura (T-IV) (Tabla 2). Posteriormente, el grupo Euroflow publicó unas guías para optimizar el diagnóstico de las diferentes neoplasias hematológicas mediante citometría de flujo<sup>84</sup>.

En 2009, se identificó un nuevo grupo de LLA-T con un perfil de expresión génica propio de etapas tempranas del proceso de diferenciación de las células T<sup>85</sup>. Este subtipo presenta un perfil inmunofenotípico concreto caracterizado por expresar CD7, pero no expresar los marcadores CD1a y CD8 y con la retención de uno o más marcadores de linaje mielóide y/o de célula madre (CD34, CD117, HLADR, CD13, CD33, CD11b, o CD65). Normalmente también expresan CD2 y CD3 citoplasmático y pueden expresar CD4, pero estos no son parte de la definición. CD5 es negativo y en los casos que sea positivo está presente en <75% de los blastos<sup>85-87</sup>. Son las denominadas leucemias linfoblásticas de precursores tempranos T (en inglés *Early T-cell precursor leukemia*, ETP-ALL) que pasaron a ser reconocidas como nuevo subtipo de LLA-T en la revisión de la OMS del año 2016<sup>88</sup>. Los pacientes ETP-ALL se caracterizan por tener una edad avanzada, un mayor porcentaje de blastos en la médula o sangre y un recuento de glóbulos blancos (en inglés, *White Blood Cells*, WBC) menor a  $50 \times 10^9/L$  en comparación con los pacientes no ETP-ALL<sup>89</sup>.

Junto a la identificación del subgrupo ETP-ALL, también se diferenció un conjunto de LLA-T con un perfil transcripcional similar a las ETP-ALL, pero con un nivel de expresión de CD5 no lo suficientemente bajo para considerarse como tal (>75% de las células). Este subgrupo se denominó *near*-ETP-ALL y presenta alteraciones genéticas diferentes a las ETP<sup>90</sup>.

**Tabla2. Principales subtipos inmunofenotípicos en la LLA-T.**

Los blastos T son TdT+ y expresan CD3+ citoplasmático (cCD3+)	
ETP-ALL	CD7+, CD1a+, CD8-, CD5+ (<75%). Al menos un marcador mielode y/o de célula madre (CD34, HLA-DR, CD13, CD33, CD117, CD11b, CD65)
Pro-T (T-I) *	CD7+, CD2-, CD5-, CD1a+, CD4+, CD8-
Pre- T (T-II) *	CD7+, CD2+ y/o CD5+, CD1a+, CD4+, CD8-
Tímica cortical (T-III) *	CD7+, CD2+, CD5+ CD1a+, sCD3 <sup>+</sup> -, CD4 <sup>+</sup> -, CD8 <sup>+</sup> -
Tímica madura (T-IV) *	CD7+, CD2+, CD5+, CD1a+, sCD3 <sup>+</sup> , CD4+ o CD8+

\*Subtipos de LLA-T propuestos por EGIL en 1995

Por último, hay un conjunto de leucemias que presentan una combinación mixta de marcadores de diferentes subtipos, son las leucemias agudas de linaje mixto (en inglés *Mixed phenotype acute leukemia*, MPAL). Se distinguen principalmente dos grupos: i) bilineales, con presencia de 2 clones con marcadores de linajes diferentes, ii) bifenotípicas, con presencia de 1 clon con mezcla de marcadores de linajes diferentes<sup>91</sup>. En relación con el componente T destacan las leucemias mixtas B/T y las leucemias mixtas de linaje mielode y linaje linfode T (T/M MPAL). Estas últimas comparten características con los subtipos de LLA-T más inmaduros y a día de hoy, para muchos casos, son difíciles de clasificar<sup>92</sup>.

### 3.5.3 Clasificación de la OMS

En 1997 la OMS propuso otro sistema de clasificación, basado en la morfología celular y en los perfiles citogenéticos, publicado finalmente en el año 2000<sup>93</sup>. Esta clasificación se revisa periódicamente. En 2008 además del estudio morfológico y citogenético, también incluyó el uso de la citometría de flujo. En la última revisión, la LLA-T se engloba dentro del grupo de proliferaciones linfoides de células T, células *Natural Killer* y linfomas<sup>75</sup>. Esta clasificación es bastante básica en cuanto a la definición de la LLA-T y solo diferencia el subtipo ETP-ALL del resto de LLA-T:

- Leucemia linfoblástica aguda T/ linfoma
- Leucemia linfoblástica aguda de precursores tempranos T / linfoma

Los autores de esta revisión indican que todavía no hay suficiente evidencia para establecer subtipos definidos genéticamente y con una relevancia clínica.

### 3.5.4 Clasificación de la ICC

Recientemente, surgió un nuevo grupo de expertos que han desarrollado la Clasificación del Consenso Internacional (en inglés, *International Consensus Classification*, ICC) de neoplasias mieloides y leucemias agudas.

Este nuevo sistema de clasificación incluye parámetros similares a la clasificación de la OMS, es decir, parámetros morfológicos, inmunofenotípicos, genéticos y moleculares<sup>94</sup>. Para la LLA-T, coincide con la última revisión de la OMS incluyendo el subtipo ETP-ALL, cuyo diagnóstico se basa en el estudio del inmunofenotipo. Además, definen un nuevo subtipo, las ETP-ALL con desregulación del factor de transcripción BCL11B<sup>95-97</sup> (ver apartado 3.5.5.1 BCL11B).

De este modo, la clasificación de la ICC para LLA-T incluye los siguientes subtipos:

- Leucemia linfoblástica aguda de precursores tempranos T con reordenamiento en *BCL11B*
- Leucemia linfoblástica aguda de precursores tempranos T
- Leucemia linfoblástica aguda T

Por último, proponen un conjunto de entidades provisionales basadas en la activación anómala de diferentes familias de factores de transcripción. El hecho de que no sean entidades definitivas es debido a que están definidas por lesiones complejas y no existe un consenso claro entre los diferentes autores que las definen. Sin embargo, existe algo común a todas estas entidades, tanto el nuevo subtipo (BCL11B) como las propuestas como provisionales, y es que generan perfiles de expresión específicos como consecuencia principalmente de alteraciones citogenéticas y moleculares (ver apartado 3.5.5.1).

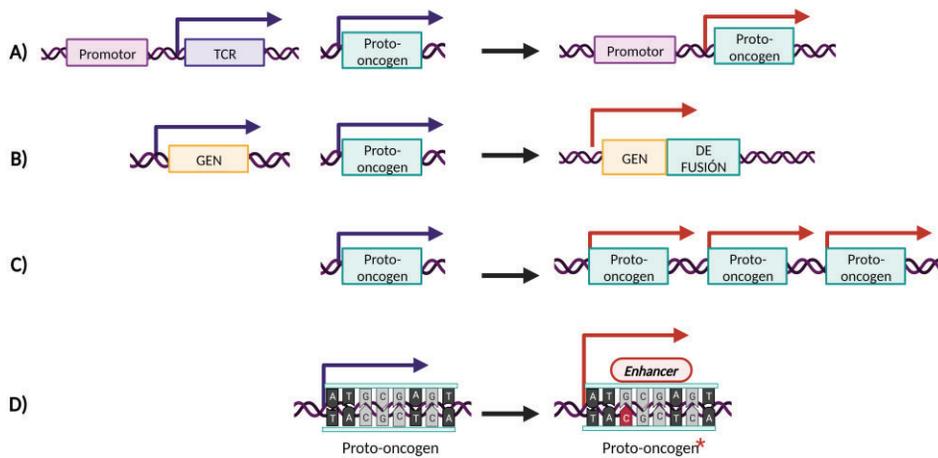
### 3.5.5 Clasificación genética

A diferencia de la LLA-B, donde el análisis de las alteraciones citogenéticas ha sido clave para su diagnóstico y clasificación, en el caso de la LLA-T no hay una clasificación citogenética unánime. Fue gracias al desarrollo de estudios de expresión (*arrays* de expresión y RT-PCR) cuando se identificó que uno de los principales eventos transformadores de la LLA-T es la expresión ectópica de factores de transcripción (FT) claves en la hematopoyesis y/o desarrollo, maduración y diferenciación de las células T. Como resultado, se genera un perfil de expresión específico que refleja el punto de bloqueo de la diferenciación del linfocito T y por tanto, el grado de inmadurez de la leucemia. Por esta razón, hay una cierta correlación con el inmunofenotipo y la firma transcriptómica de los blastos T.

Este tipo de alteraciones se denominan alteraciones de tipo A o primarias, son mutuamente excluyentes, se pueden dar sin necesidad de que haya un reordenamiento cromosómico y permiten clasificar a un 70% de las LLA-T, aproximadamente<sup>98</sup>.

La activación anómala de estos factores de transcripción puede deberse a varios defectos genéticos y epigenéticos (Fig. 4):

1. Traslocaciones cromosómicas que involucran el gen receptor de las células T (TCR alfa, TCRA; TCR delta, TCRD; TCR beta, TCRB). Afectan a entre el 35-50% de los pacientes con LLA-T<sup>99</sup> (Fig. 4A).
2. Reordenamientos cromosómicos que llevan a la formación de genes de fusión con actividad oncogénica (Fig. 4B).
3. Duplicaciones/amplificaciones de factores de transcripción con actividad oncogénica (Fig. 4C).
4. Mutaciones o pequeñas inserciones que generan nuevas secuencias reguladoras que actúan como activadoras de promotores de oncogenes (Fig. 4D).



**Figura 4. Alteraciones que dan lugar a la activación de factores de transcripción definitorios de subtipos de LLA-T**

Junto a las alteraciones primarias, durante el desarrollo de la LLA-T, se suceden otro conjunto de anomalías que afectan a genes y vías de señalización, provocando la alteración de procesos de diferenciación celular, apoptosis, activación de oncogenes e inhibición de genes supresores de tumores; llevando todo ello, a una proliferación de las células neoplásicas. Son las denominadas alteraciones de tipo B o alteraciones secundarias, que se distribuyen por los diferentes subgrupos definidos según la expresión de FT, formando de este modo combinaciones de alteraciones que definen vías leucemogénicas únicas. Tanto las alteraciones tipo A como tipo B se consideran alteraciones “*drivers*” de la LLA-T, es decir, confieren a las células transformadas una ventaja que les permite proliferar y evolucionar a lo largo del desarrollo de la leucemia.

#### 3.5.5.1 Subtipos genéticos según factor de transcripción afectado

Los factores de transcripción (FT) responsables de generar estos perfiles de expresión específicos son: BCL11B, factores incluidos en la familia de genes *homeobox* (HOXA, TLX1 (HOX11), TLX3 (HOX11L2), NKX2), factores incluidos en la familia de genes *basic helix-loop-helix* (TAL1, TAL2, LYL1 y BHLHB1 (OLIG2)), LMO1 y LMO2:

- **BCL11B**

*BCL11B* codifica para un FT esencial durante las primeras etapas del proceso de diferenciación de las células T. En la LLA-T, este gen tiene un papel diferente en función del tipo de alteración que presenta. Por un lado, actúa como gen supresor de tumores, viéndose afectado por mutaciones y deleciones (ver apartado 3.5.2.2). Por otro, tiene función oncogénica como consecuencia de traslocaciones/inserciones que lo sitúan próximo a diversos *enhancers*, y llevan a su activación transcripcional, dando lugar a un perfil de expresión específico que identifica un nuevo subtipo de leucemia, descrito tanto en pacientes pediátricos como adultos<sup>95,96</sup>.

Este nuevo subtipo genético engloba tanto leucemias LLA de tipo ETP-ALL como LMA inmaduras. Más del 80% de los casos presentan mutaciones activadoras de *FLT3*<sup>95,96</sup>.

En cuanto al pronóstico, con los pocos casos que hay hasta ahora, parece que los pacientes con este tipo de reordenamiento tienen un mejor pronóstico que los pacientes con una LLA-T de tipo ETP sin reordenamientos en *BCL11B*<sup>95</sup>.

- **Familia *Homeobox* (HOXA, TLX1, TLX3, NKX2):**

- ***HOXA***

El *cluster* de genes *HOXA* es uno de los 4 *clusters* en los que se organizan los genes de la familia *homeobox* (*HOX*). Cada *cluster* está formado por entre 9 a 11 genes con una estructural conservada, una expresión estrictamente regulada<sup>100-103</sup> y están implicados en el correcto desarrollo embriológico, como también en la regulación de la hematopoyesis y leucemogénesis<sup>104,105</sup>.

Desregulaciones en la expresión de los genes *HOXA9* y *HOXA10* se han descrito por estar implicadas en la patogénesis de la LLA-T en un 3% de casos<sup>106</sup>. Estos cambios de expresión pueden ser debidos a mecanismos *cis*, como son las traslocaciones cromosómicas con los genes *TCRB* y *TCRG*, mayoritariamente<sup>107</sup> y las inversiones en el cromosoma 7 (*inv(7)(p15q34)*), donde se encuentran los genes *HOXA*, que también sitúan al gen próximo al *TCRB* produciéndose una sobre-expresión<sup>108</sup>.

Además, la expresión de los genes *HOXA* también puede estar desregulada por la expresión de otros oncogenes activados por fusiones como son: *KMT2A::ENL*<sup>109</sup>, *PICALM::MLL10*<sup>107,110</sup> y *SET::NUP214*<sup>111</sup>; siendo estos mecanismos de activación en *trans*.

Las LLA-T con este perfil de expresión presentan un inmunofenotipo inmaduro, ETP o *near-ETP* en un 40-45% de los casos<sup>112,113</sup>.

Varios grupos han asociado este subtipo a una mala supervivencia, como también las alteraciones genéticas que se asocian a este perfil de expresión<sup>113-116</sup>. No obstante, no se han visto los mismos resultados en los casos en los que el inmunofenotipo asociado no es una ETP-ALL<sup>113,115,117</sup>.

- ***TLX1***

Hox11 (*TLX1*), junto con Hox11L1 (*TLX2*) y Hox11L2 (*TLX3*) pertenecen a una familia de genes *homeobox* no incluida en los cuatro *clusters* de genes *homeobox* que se comentaba en el apartado anterior<sup>118</sup>.

El gen *TLX1* se identificó estudiando los puntos de rotura de las traslocaciones *t(10;14)(q24;q11)* que se dan en las leucemias con dicho reordenamiento<sup>119</sup>. Esta traslocación cromosómica aparece en un 5-10% de los casos pediátricos y un 30% de las LLA-T en adultos<sup>109,120</sup>. Tiene como consecuencia la aproximación del gen *TLX1* a los genes *TCRA* y *TCRD* quedando bajo el control de sus promotores lo que lleva a un incremento de los niveles de expresión de *TLX1*<sup>119</sup>.

Estudios realizados en ratones mostraron que los cambios en la expresión de *TLX1* llevaban a un bloqueo de la diferenciación de los timocitos en la etapa cortical<sup>121,122</sup>.

Se trata de un subtipo con buen pronóstico. Esto se debe principalmente a que los linfoblastos TLX1+ se caracterizan por expresar genes asociados a proliferación celular y por tener una menor expresión de *BLC2* y otros genes anti-apoptóticos. Por tanto, estos linfoblastos responden de manera favorable a fármacos que inducen una muerte celular programada<sup>109</sup>.

### ○ **TLX3**

Se trata de otro de los genes *TLX*. Similar a *TLX1*, *TLX3* (*HOX11L2*) se identificó estudiando los puntos de rotura de otros de los reordenamientos frecuentes de la LLA-T, en concreto la traslocación t(5;14)(q35;q32)<sup>123</sup>, presente en un 20-25% de las LLA-T pediátricas y un 5% de las LLA-T adultas<sup>109,120,123-125</sup>. Esta traslocación deja al gen *TLX3* bajo el control del gen *BCL11B*<sup>123</sup>. No obstante, en un número menor de casos, también se han identificado reordenamientos de *TLX3* con *TCRB*, *TCRD*, *CDK6* u otros genes desconocidos<sup>126,127</sup>.

Cabe decir que el perfil de expresión que presentan las leucemias de tipo TLX1 y TLX3 es muy similar y, al igual que lo que ocurre con en el subtipo TLX1, en el caso de TLX3 también se produce un bloqueo de la diferenciación de los timocitos en la etapa cortical, en la mayoría de los casos (40%). No obstante, también hay LLA-T TLX3+ de tipo inmaduras y maduras<sup>126</sup>.

En cuanto al pronóstico difiere del subtipo TLX1, pues en pacientes pediátricos se ha asociado con bueno y mal pronóstico o no se ha encontrado una asociación<sup>114,124,128,129</sup>. Estos datos contradictorios podrían deberse al número diferente de LLA-T de tipo inmaduro y cortical incluidas en las cohortes de estudio.

### ○ **NKX2**

La familia de genes *NKX2* también pertenecen a los genes tipo *homeobox*. Esta familia se ha asociado históricamente al desarrollo de la LLA-T. El gen *NKX2-5* se identificó como parte de un reordenamiento oncogénico con el gen *TCR $\alpha$* <sup>30</sup> y con el gen *BCL11B*<sup>131</sup>. Por su parte, el gen *NKX3-1* se halló altamente expresado en los casos con reordenamiento TAL1<sup>107</sup>. Fue en el trabajo realizado por Homminga *et al.*<sup>132</sup> cuando identificaron una expresión aberrante en los genes *NKX2-1* y *NKX2-2* debido a traslocaciones en las que participaba el gen *TCR* o a inversiones cromosómicas. Estos últimos mecanismos resultan en perfiles de expresión específicos que caracterizan al 5% de LLA-T pediátricas<sup>132</sup>.

Los perfiles de expresión de los genes *NKX2-1* y *NKX2-2* agrupan en el mismo *cluster* a las muestras con estos reordenamientos. A su vez, los perfiles de expresión de este grupo son similares a los del subtipo TLX1 y a los de los timocitos en la etapa temprana cortical<sup>132</sup>. De hecho, en el trabajo de Homminga *et al.*<sup>132</sup> consideran este subtipo junto al subtipo TLX1 dentro de un mismo grupo al que denominan *cluster* proliferativo. En cuanto al pronóstico de las LLA-T de tipo NKX2 es similar al del subtipo TLX1.

En cambio, la expresión aberrante de *NKX3-1* está asociada al subgrupo TAL/LMO<sup>107</sup> y las traslocaciones que afectan al gen *NKX2-5* están asociadas a una LLA-T más inmadura<sup>132</sup>.

- **Familia BHLH (TAL1, TAL2, LYL1)**

- **TAL1**

*SCL/TAL1 (Stem cell leukemia/T cell acute leukemia 1)* es un gen perteneciente a la familia de factores de transcripción BHLH (*basic helix-loop-helix*) que se expresa predominantemente en megacariocitos y eritrocitos y es esencial para el desarrollo de la hematopoyesis. Durante el desarrollo de los timocitos es reprimido mediante mecanismos epigenéticos<sup>133-135</sup>. No obstante, aunque *TAL1* no se expresa normalmente en células linfoides, sí que se observa una expresión de este factor en el 30-45% de las LLA-T pediátricas y 10-15% de las LLA-T adultas<sup>109,120</sup>.

Hay diferentes mecanismos que pueden llevar a la expresión anómala de *TAL1*: i) microdelección que yuxtapone el gen *TAL1* a la región promotora de su gen vecino *STIL*, formándose el gen de fusión *STIL::TAL1* que lleva a un incremento de la expresión de *TAL1*<sup>136-138</sup> y que se da en un 25% de las LLA-T; ii) traslocación t(1;14)(p32;q11) que lleva a *TAL1* a ser regulado por los genes *TCRA* o *TCRD*<sup>39,140</sup> y que aparece en el 3% de las LLA-T pediátricas; iii) mutaciones somáticas en una región intergénica anterior al extremo 5' de *TAL1* que crean un sitio de unión para *MYB* el cual actúa como *enhancer*<sup>141,142</sup>; iv) alteraciones epigenéticas que afectan al plegamiento de la cromatina y que provocan una sobre-expresión de *TAL1* vía *enhancers*<sup>143-145</sup>.

En las LLA-T que presentan un perfil de expresión TAL1+, el estadio de maduración en el cual se ha producido el bloqueo del ciclo celular de los timocitos se corresponde con una etapa cortical tardía o madura, caracterizándose por la expresión del marcador de superficie CD3+ y los marcadores CD4+ y CD8+<sup>146</sup>.

Los pacientes con este perfil de expresión tienen un pronóstico poco favorable debido principalmente a la sobre-regulación de moléculas anti-apoptóticas<sup>147</sup> lo que haría que no respondieran bien a las terapias convencionales.

- **TAL2**

A diferencia de *TAL1*, el gen *TAL2 (T cell acute leukemia 2)*, se encuentra alterado en muy pocos casos de LLA-T. Este gen fue identificado por estar implicado en la traslocación t(7;9)(q34;q32), que tendría como consecuencia su yuxtaposición junto al *TCRB*<sup>148</sup>. Posteriormente, se identificó la sobre-expresión de *TAL2* en un caso de LLA-T, aunque no se encontraron alteraciones citogenéticas asociadas a esta expresión aberrante<sup>146</sup>.

- **LYL1 y BHLHB1**

Los genes que codifican a los factores LYL1 y BHLHB1 (conocido como OLIG2) también se encuentran expresados de manera anómala como consecuencia de traslocaciones cromosómicas que los sitúan próximos a los genes del *TCR*. No obstante, se trata de alteraciones que aparecen en muy pocos casos de LLA-T<sup>140,149,150</sup>.

- **Familia LIM**

*LMO1* y *LMO2* pertenecen a la familia de genes *LIM domain only* (LMO) y se encuentran sobre-expresados de manera patogénica en aproximadamente el 15-20% de las LLA-T.

En un 10% de los casos, el incremento de la expresión es debido a traslocaciones cromosómicas con el gen *TCRA* en el caso de *LMO1*<sup>151</sup>, y con el *TCRD* en el caso de *LMO2*<sup>152</sup>. Otro mecanismo de activación es por la presencia de mutaciones en sus regiones promotoras que crean nuevos sitios de unión para genes que actúan como *enhancers*, y que se dan en un 2% de las LLA-T en el caso de *LMO1*<sup>153</sup> y entre el 3-5% en el caso de *LMO2*<sup>154</sup>. Finalmente *LMO2* también puede activarse por pequeñas deleciones que se dan en regiones cromosómicas cercanas a este gen y que aparecen en hasta un 5% de LLA-T<sup>155,156</sup>.

Las proteínas LMO forman complejos transcripcionales con *TAL1* y otros factores de la familia *bHLH*<sup>157</sup>. De hecho, es frecuente la co-expresión de *LMO1* y *LMO2* junto a *TAL1*<sup>146</sup> y la actividad oncogénica de estos factores de transcripción es potenciada por el efecto de *TAL1* en modelos de ratón transgénico<sup>157-159</sup>.

Dado que estos genes suelen estar alterados junto a *TAL1*, se suelen agrupar en una misma categoría TAL/LMO, pues tienen un perfil de expresión similar y se relacionan con un inmunofenotipo cortical tardío/ maduro donde ya se expresa el marcador de superficie CD3+. No obstante, también es frecuente la co-expresión de *LMO2* junto a *LYL1*, ambos se encuentran sobre-expresados en LLA-T más inmaduras y en el subtipo ETP-ALL<sup>146,160</sup>.

En cualquier caso, la mayoría de las LLA-T con *LMO1/LMO2* activado se caracterizan por tener marcadores de mal pronóstico: baja incidencia de mutaciones en *NOTCH1*, alteraciones en *PTEN*, deleciones en el 6q, traslocaciones en *MYC*, sobre-expresión de *MEF2C*<sup>126,132</sup>.

- **MEF2C**

En el trabajo de Homminga *et al.*<sup>132</sup> se describió un *cluster* inmaduro de LLA-T sin alteraciones previamente descritas y que incluía un 30% de casos con altos niveles de expresión de *MEF2C*.

Experimentos funcionales en líneas celulares de LLA-T demostraron el papel de *MEF2C* como regulador transcripcional de genes que están altamente expresados en LLA-T inmaduras, como *LMO2*, *LYL1* y *HHEX*. Además, también se observó que la expresión de *MEF2C* provocaba un bloqueo de la diferenciación de las células T en estadios inmaduros<sup>132</sup>. No obstante, también se han encontrado altos niveles de expresión de *MEF2C* en LLA-T no ETP<sup>161,162</sup>.

La expresión de *MEF2C* se ha asociado a una peor respuesta a glucocorticoides<sup>161</sup>.

### 3.5.5.2 Alteraciones genéticas secundarias

Las alteraciones tipo B o secundarias se dividen en diferentes bloques según la función de los genes afectados.

- **Factores de transcripción con función de oncogenes**

Hay factores de transcripción que actúan como oncogenes y que no definen un subtipo específico de LLA-T:

- **MYC**

*MYC* codifica para un FT que es un regulador esencial del desarrollo de las células T. Concretamente, durante la fase temprana del desarrollo se encarga de regular el control del crecimiento celular, siendo un factor controlado por la vía de señalización de NOTCH1 y la señalización del pre-TCR<sup>163</sup>. La sobre-expresión de *MYC* afecta a aproximadamente un 6% de pacientes adultos y pediátricos<sup>164</sup>. Solo en el 1% de las LLA-T esta sobre-expresión es consecuencia de una traslocación con el *TCRA* o el *TCRD*<sup>63,165,166</sup>, y en la mayoría de los casos (5%) viene dada por la creación de un *enhancer* regulado por *NOTCH1* (N-Me, *NOTCH1-controlled MYC enhancer*)<sup>167,168</sup>. Al igual que *NOTCH1*, *MYC* también es una diana de *FBXW7* para ser degradado vía proteosoma y por tanto, las mutaciones en *FBXW7* producen un incremento de los niveles de las proteínas de *MYC*<sup>169,170</sup>.

La sobre-expresión de *MYC* es frecuente en el subtipo TAL/LMO1<sup>171</sup> y hay estudios que han descrito esta sobre-expresión como marcador de alto riesgo, pues se ha asociado a la presencia de hiperleucocitosis, peor respuesta a los glucocorticoides y a la quimioterapia estándar<sup>171-173</sup>.

- **MYB**

*MYB* codifica para un FT que regula múltiples procesos celulares como la proliferación, diferenciación, ciclo celular, apoptosis, angiogénesis y adhesión celular<sup>174</sup>. La activación de *MYB* puede ocurrir a través de la traslocación t(6;7)(q23;q34) que se da en el 1% de los pacientes con LLA-T, siendo generalmente niños de menos de 2 años<sup>175,176</sup>. También se han descrito duplicaciones focales del gen que llevan a un incremento de su expresión en un 10% de pacientes, tanto niños como adultos<sup>177,178</sup>. Además, *MYB* también puede ser regulado por otros mecanismos como la activación transcripcional de *TAL1*<sup>179</sup> o mediante la infra-regulación (*downregulation*) de microARN que tienen como diana este oncogen<sup>180,181</sup>.

Las alteraciones que afectan al gen *MYB* no aparecen en ningún paciente con el subtipo ETP-ALL<sup>176</sup> y se han asociado a un bloqueo de la diferenciación en la fase cortical<sup>132</sup>. Por tanto son frecuentes en los grupos HOXA, TLX1/NKX2-1 y TLX3 positivos<sup>176</sup>.

- **MLL**

*MLL* (*KMT2A*) es un gen responsable de regular otros genes que están involucrados en el proceso hematopoyético. Este gen fue descrito en pacientes pediátricos de LMA, pero también está involucrado en la patogénesis de la LLA-T. Se han descrito más de 135 fusiones diferentes en las que participa el gen *MLL*, en los distintos tipos leucemias<sup>182</sup>. En el caso de

la LLA-T, el análisis en una cohorte pediátrica de 100 pacientes identificó reordenamientos en 12 pacientes que incluyeron: *MLL::AF6 (KMT2A::MLLT4)*, del3'*MLL*, *MLL::ENL (KMT2A::MLLT1)*, *MLL::PICALM (KMT2A::PICALM)* y *MLL::AF17 (KMT2A::MLLT6)*<sup>117,183</sup>.

El pronóstico de las LLA-T con reordenamiento en *MLL* es generalmente desfavorable, con la excepción del reordenamiento *MLL::ENL* que se ha visto que presenta buen pronóstico<sup>184</sup>.

- **Factores de transcripción con función variable (gen supresor/oncogén)**

La presencia de mutaciones somáticas y deleciones llevan a la disrupción de factores de transcripción que tienen función de supresores tumorales. En condiciones determinadas, estos genes también pueden actuar como oncogenes. A continuación, se detallan estas funciones:

- ***RUNX1***

*RUNX1* codifica a un FT clave en el desarrollo de la hematopoyesis y en la diferenciación de los linfocitos T<sup>185</sup>. Está mutado aproximadamente en un 16% de los pacientes adultos<sup>186,187</sup> y un 9% de los pacientes pediátricos<sup>188,189</sup>, principalmente del subtipo ETP-ALL<sup>186-190</sup>. Estudios publicados por un grupo alemán (GMALL) asociaron las mutaciones en este gen a una peor supervivencia<sup>186,187</sup>.

En la LLA-T, la mayoría de las mutaciones descritas llevan a una pérdida de función del gen<sup>186-190</sup>, lo que sugiere un papel como supresor tumoral. En cambio, hay estudios que han mostrado que la proteína *RUNX1* coopera con *TAL1* en el desarrollo de la LLA-T<sup>191</sup> y que su expresión es necesaria para mantener la supervivencia de las células LLA-T, conjuntamente con *TAL1* y *NOTCH1*<sup>192</sup>, sugiriendo un papel oncogénico.

Dado que la mayoría de los pacientes con mutaciones en *RUNX1* pertenecen al subtipo ETP-ALL<sup>186-188,190</sup>, donde no son comunes las alteraciones en *NOTCH1* y *TAL1*, se sugiere que en este subtipo *RUNX1* podría ejercer como supresor tumoral mientras que en los casos no ETP-ALL una función más oncogénica<sup>193</sup>.

Se han descrito mutaciones en *RUNX1* en la línea germinal de familias con trastorno plaquetario familiar con neoplasia mieloide asociada (en inglés *Familial Platelet Disorder with Associated Myeloid Malignancies*, FPDMM) donde se observó una mayor incidencia de LLA-T, sugiriendo una relación de estas mutaciones germinales con la predisposición a la LLA-T<sup>194</sup>.

- ***ETV6***

El gen *ETV6* es un represor transcripcional esencial en el desarrollo de las células madre hematopoyéticas<sup>195</sup>. Mutaciones y deleciones en este gen están presentes en el 8% de pacientes pediátricos y en el 14% de pacientes adultos<sup>80,196</sup> y son especialmente frecuentes en el subtipo ETP-ALL, donde afectan a un 24% de los pacientes<sup>197</sup>. Ensayos funcionales han demostrado que estas mutaciones llevan a una pérdida de función de este gen supresor de tumores<sup>197</sup>.

### ○ **GATA3**

*GATA3* codifica para un regulador transcripcional expresado solo en el linaje T y es esencial durante las etapas tempranas del proceso de diferenciación de los linfocitos T<sup>198</sup>. La alteración más frecuente es una mutación puntual que lleva a la pérdida de función del gen y que afecta su dominio de unión al ADN<sup>199</sup>. Se han descrito mutaciones en este gen en el 5% de los pacientes pediátricos y en un 3% de adultos<sup>196</sup>, siendo una mutación más frecuente en el subtipo ETP-ALL<sup>189,200</sup>. En este subtipo *GATA3* puede estar metilado en un 33% de los pacientes, llevando a una disminución de la expresión del mismo<sup>200</sup>.

No obstante, también hay trabajos que demuestran que *GATA3* puede actuar como oncogén<sup>201</sup>. Aunque se conoce poco acerca de los mecanismos de acción de estas alteraciones, se ha descrito que *GATA3* interviene regulando la accesibilidad de la cromatina y modula la actividad de un *enhancer* de *MYC*, siendo necesario para el inicio y mantenimiento de la LLA-T<sup>202</sup>.

Estos datos muestran que *GATA3* también puede tener diferentes funciones en el desarrollo de la LLA-T. Presenta una menor expresión en el subtipo ETP-ALL, coherente con el hecho de que su función sea promover la diferenciación de los linfocitos T y, por tanto, un silenciamiento de este podría llevar a un bloqueo en etapas más tempranas favoreciendo la transformación leucémica. Mientras que tendría un papel más oncogénico en las etapas más tardías de la LLA-T.

### ○ **BCL11B**

Se trata de un FT implicado en la diferenciación de los linfocitos T mediante la represión de la diferenciación hacia los linajes mieloides y NK, favoreciendo la del linaje T<sup>203</sup>. Estudios en ratones han mostrado que la inactivación de *BCL11B* lleva a una parada de la diferenciación de los linfocitos T en una etapa cortical<sup>204,205</sup> y a una actividad de autorrenovación aberrante<sup>206</sup>. En humanos, se encuentra mutado o deleciónado en un 10% de casos pediátricos<sup>207,208</sup> y un 9% de adultos<sup>196</sup>, lo que sugiere una función como supresor tumoral. Las alteraciones en este gen son especialmente frecuentes en los subtipos TLX1 y TLX3<sup>209</sup>.

No obstante, como se ha explicado en un apartado 3.5.5.1, *BCL11B* también está implicado en alteraciones que llevan a un incremento de su expresión y que definen subtipos de LLA-T más inmaduras. Además, se han descrito traslocaciones donde interviene este gen que llevan a la activación de *TLX3* o *NKX2-5*<sup>22</sup>.

### ○ **LEF1**

*LEF1* codifica para un FT, perteneciente a la familia de factores LEF/TCF, implicado en la vía de señalización WNT<sup>210</sup>. Se expresa durante el desarrollo de las células T, especialmente en la etapa cortical (CD1a<sup>+</sup>)<sup>30</sup> y la pérdida de su expresión contribuye a una parada del proceso de diferenciación de las células T<sup>211,212</sup>. Se han descrito deleciones mono-alélicas y bi-alélicas y mutaciones que afectan a *LEF1* en un 10-15% de pacientes pediátricos<sup>213</sup> y un 2% de pacientes adultos<sup>196</sup>, sugiriendo la necesidad de su inactivación para la progresión de la LLA-T.

Se desconoce su implicación en el desarrollo de la LLA-T, pero hay evidencias que sugieren que su inactivación podría combinarse junto a la activación de *NOTCH1* promoviendo una sobre-regulación de la expresión de *MYC*<sup>214-216</sup>. De hecho, pacientes con inactivación de *LEF1* muestran altos niveles de expresión de *MYC*<sup>213,217</sup>. Estos pacientes se caracterizan por tener un perfil de expresión propio de etapas corticales tempranas, similar al perfil de las LLA-T TLX1+<sup>213</sup>.

No obstante, también se ha descrito que *LEF1* puede tener un papel como oncogén en el desarrollo de la LLA-T, y su función parece estar determinada por el momento del proceso leucemogénico en el que se deleciona. Si *LEF1* está presente durante la transformación y por tanto actuando como oncogén, la leucemia crea una dependencia al mismo. En cambio, si *LEF1* está delecionado, el desarrollo de las células T se ve alterado y se favorece el desarrollo de la leucemia, como se ha explicado<sup>218</sup>.

- ***WT1***

Este FT está implicado en el mantenimiento del estado quiescente de las células madre primitivas y en promover la diferenciación de progenitores más maduros<sup>219</sup>. Mutaciones y delecciones en este gen se encuentran en aproximadamente el 10% de los pacientes pediátricos y adultos de LLA-T<sup>196,220</sup>, llevan a una pérdida de función y están asociadas con los subtipos de LLA-T con expresión de los oncogenes TLX1, TLX3 y HOXA<sup>220</sup>. Con la identificación del nuevo subtipo de LLA-T caracterizado por la activación de BCL11B, se observó que las mutaciones en *WT1* eran frecuentes en este subtipo<sup>95,96</sup>. Además, las alteraciones bi-alélicas en *WT1* son una característica de las leucemias mixtas T/M MPAL<sup>221</sup>

Las alteraciones que afectan a este gen tienen mayor prevalencia en pacientes recaídos<sup>222</sup> y se han asociado con una peor supervivencia libre de recaída en una cohorte de pacientes adultos (GMALL)<sup>223</sup>.

- **Alteraciones en vías de señalización**

- **Vía de señalización *NOTCH1*:**

El gen *NOTCH1* codifica para una glicoproteína de transmembrana de tipo I. Tras su activación, mediante la unión de un ligando, actúa como FT y transforma señales extracelulares en cambios transcripcionales<sup>224</sup>. En el sistema hematopoyético, la activación de *NOTCH1* en el timo es crucial para la especificación temprana de las células T y el desarrollo de los timocitos<sup>13</sup>. Concretamente, está involucrado en la diferenciación de los linfocitos en la etapa cortical<sup>225</sup>.

En la LLA-T, *NOTCH1* actúa como oncogén induciendo el desarrollo de la leucemia<sup>226</sup>. Esta función oncogénica se dedujo por el rápido desarrollo de la neoplasia en ratones trasplantados con progenitores hematopoyéticos infectados con retrovirus que llevaban a la expresión constitutiva de *NOTCH1*<sup>227</sup>. Análisis posteriores de sus genes diana y de los perfiles de expresión resultantes, tras la activación de este gen, mostraron que podía promover el desarrollo de la leucemia a través de una *up-regulación* transcripcional de vías de señalización anabólicas como la biosíntesis de ribosomas, la traducción de proteínas y el

metabolismo de nucleótidos y aminoácidos<sup>228,229</sup>. La actividad desencadenada por el programa transcripcional de *NOTCH1* puede ser potenciada por el efecto de la *up-regulación* del oncogén *MYC*, uno sus principales genes diana<sup>167,168,228</sup>.

La primera evidencia de la presencia de alteraciones genéticas en *NOTCH1* en la LLA-T fue por la identificación de traslocaciones poco comunes con el gen *TCRB*<sup>230</sup> y *TCRD*<sup>231</sup> que llevaba a su activación<sup>230</sup>. Sin embargo, el principal motivo de la activación de la señalización de *NOTCH1* es por la presencia de mutaciones de ganancia de función que se dan principalmente en los dominios HD y PEST<sup>232</sup>. Las mutaciones en el dominio HD llevan a una activación del receptor sin necesidad de unión de su ligando, mientras que las mutaciones en el dominio PEST producen un aumento de la estabilidad de la región intracelular del receptor (en inglés *intracellular domain of Notch*, ICN1) y una activación prolongada de NOTCH1<sup>232</sup>. Todas estas mutaciones llevan a un incremento de la actividad de la vía de señalización de NOTCH1 desencadenándose una sobre-regulación de sus vías efectoras (PI3K/Akt/mTOR<sup>233</sup>, c-myc<sup>228</sup>, y NF-kb<sup>234</sup>, principalmente) implicadas en el control transcripcional del crecimiento celular y metabolismo.

Se han descrito mutaciones que afectan a *NOTCH1* en el 60% de los pacientes pediátricos<sup>232</sup>, mientras que en los pacientes adultos esta frecuencia varía según la cohorte estudiada siendo superior en todos los casos al 50%<sup>235-237</sup>. Estas alteraciones se encuentran en el 70% de los subtipos TLX1+, TLX3+, y NKX2-1+<sup>112</sup>; es decir, subtipos relacionados con el bloqueo de la diferenciación en la etapa cortical, ya que es en esta etapa del desarrollo en la que principalmente interviene la vía NOTCH1. En cambio, su incidencia disminuye tanto en las LLA-T más inmaduras como en las más maduras, siendo menor al 40%<sup>189,238</sup>.

Las mutaciones de *NOTCH1* pueden ser eventos primarios tempranos, incluso se ha descrito que pueden surgir en un estadio prenatal<sup>239</sup>, mientras que en un estudio de *single cell* se ha mostrado que estas mutaciones tienden a ocurrir en estadios tardíos del desarrollo leucemogénico. El hecho de que estas mutaciones aparezcan tanto como eventos tempranos como tardíos, sugiere que en algunos casos puedan ser eventos secundarios de la LLA-T<sup>240</sup>.

Las mutaciones de pérdida de función en el gen *FBXW7*, que codifica para una subunidad del complejo proteico ubiquitina ligasa E3<sup>241</sup>, son otro mecanismo responsable de la activación de la señalización de NOTCH1<sup>241,242</sup>. La repercusión de estas mutaciones es similar a lo que ocurre con las mutaciones del dominio PEST de NOTCH1, pues se produce un incremento de la estabilidad de la forma activa de NOTCH1, ICN1. Además, las mutaciones en *FBXW7* también afectan a la degradación de otras oncoproteínas<sup>243-245</sup>. En aproximadamente el 25% de los pacientes con LLA-T aparecen mutaciones en el dominio HD de *NOTCH1* junto a mutaciones en el dominio PEST o en el gen *FBXW7*, lo que da como resultado una activación doble de *NOTCH1*, combinando la activación independiente del ligando y la estabilidad prolongada de ICN1<sup>232</sup>.

Dado que *NOTCH1* es el gen más mutado en la LLA-T existen múltiples estudios donde analizan su impacto pronóstico, de manera aislada o junto al gen *FBXW7*. En un primer estudio, donde se analizaron 157 pacientes pediátricos, tratados de acuerdo con el protocolo multicéntrico ALL-BFM 2000, se mostró que las mutaciones en *NOTCH1* se correlacionaban

con una buena respuesta a la prednisona, una ERM favorable y un buen pronóstico a largo plazo<sup>246</sup>. Lo mismo ocurrió en el análisis de 55 pacientes pediátricos, publicado por un grupo de trabajo japonés, en el que aquellos pacientes con mutaciones en *NOTCH1* y *FBXW7* mostraron un pronóstico favorable<sup>247</sup>. En otro estudio, en este caso 141 pacientes adultos de la cohorte francesa (GRAALL), las mutaciones de *NOTCH1* y *FBXW7* se asociaron con un pronóstico favorable<sup>237</sup>. Sin embargo, en la cohorte inglesa (UKALL), compuesta por pacientes adultos, no hubo asociación entre las mutaciones en *NOTCH1* y *FBXW7* y el pronóstico<sup>248</sup>. En otros estudios en pacientes pediátricos, tratados de acuerdo con protocolos DCOG, aunque las mutaciones en *NOTCH1* y *FBXW7* se asociaron con una mejor respuesta al tratamiento, esto no supuso mejoras a largo plazo<sup>249,250</sup>. Los resultados de la cohorte francesa (GRAALL), compuesta por pacientes adultos, mostraron que el impacto favorable de las mutaciones en *NOTCH1/FBXW7* dependía del protocolo de tratamiento utilizado<sup>251</sup>. Por tanto, no hay resultados concluyentes sobre el pronóstico de estos genes.

○ **Vía de señalización PI3-AKT-mTOR:**

Esta vía de señalización está implicada en el control de multitud de procesos celulares como la proliferación, supervivencia, crecimiento, metabolismo y angiogénesis<sup>252</sup>; y es otra de las vías más frecuentemente activadas en la LLA-T. Su activación aberrante se debe a alteraciones en los genes que codifican para sus diferentes miembros, siendo las más frecuentes las que afectan a *PTEN*, un regulador negativo de la vía<sup>253</sup>.

Diferentes estudios han mostrado que la pérdida de la actividad de *PTEN* en células hematopoyéticas progenitoras lleva a un desarrollo de la LLA-T<sup>254-256</sup>. En particular, se ha observado que la delección de *PTEN* en HSC de hígado fetal de ratones conduce a desarrollar una LLA-T<sup>257</sup>.

Los principales mecanismos que llevan a la inactivación de *PTEN* son delecciones, mutaciones de pérdida de función y microdelecciones originadas por eventos de recombinación<sup>258</sup>. Otro mecanismo menos frecuente, observado en un 1% de pacientes, es una delección de entre 200 kb-1,4 Mb que afecta a una región *enhancer* de *PTEN*<sup>259</sup>. En total, la frecuencia de las alteraciones que afectan a este gen es variable en las diferentes cohortes de estudio, siendo de entre 11-27% en el caso de pacientes pediátricos<sup>209,258,260-262</sup> y entre 5-17% en el caso de los pacientes adultos<sup>263,264</sup>.

En pacientes pediátricos, estas alteraciones se identifican principalmente en el subtipo TAL/LMO<sup>209,261</sup> mientras que en adultos raramente se dan en pacientes que tienen mutaciones en *NOTCH1/FBXW7* y son frecuentes en pacientes STIL/TAL1 positivos<sup>263</sup>.

En general las alteraciones en *PTEN* se han asociado a peor respuesta a la quimioterapia, mayor probabilidad de recaer y un mal pronóstico a largo plazo, tanto en cohortes pediátricas como adultas<sup>260-270</sup>. Un estudio en pacientes pediátricos mostró que el mal pronóstico de las mutaciones inactivadoras de *PTEN* podía ser compensado por la presencia de mutaciones activadoras en *NOTCH1*<sup>271</sup>. Otro estudio, que incluyó pacientes pediátricos y adultos, concluyó que el impacto pronóstico de la pérdida de *PTEN* dependía del tipo de alteración, pues mientras que las delecciones estuvieron relacionadas con supervivencias más bajas,

tanto en niños como en adultos, las mutaciones no se asociaron a un mal pronóstico<sup>264</sup>. No obstante, en un estudio posterior se mostró que todas las alteraciones que afectan a *P TEN* estaban asociadas a un pronóstico desfavorable en una cohorte de pacientes pediátricos<sup>270</sup>. Por lo tanto, el impacto clínico de las alteraciones de *P TEN* sigue siendo un gran debate.

Otras mutaciones que afectan a la vía de señalización PI3K–AKT incluyen mutaciones activadoras de *AKT1* y mutaciones en las subunidades catalíticas y reguladoras de *PI3K*, todas ellas detectadas en baja frecuencia en la LLA-T<sup>260</sup>.

#### ○ **Vía de señalización IL7-JAK-STAT**

La vía de señalización activada por la IL-7 tras su unión al receptor IL-7R es necesaria para el proceso de diferenciación, proliferación y supervivencia de las células progenitoras tempranas T<sup>272</sup>. Esta activación induce la fosforilación recíproca de JAK3 y JAK1 que a su vez lleva al reclutamiento y activación del regulador transcripcional STAT5. Este factor homodimeriza y migra al núcleo para actuar sobre sus genes diana, entre los que destacan genes reguladores de la apoptosis (*BCL-2*, *BCL-XL* y *MLC1*)<sup>272,273</sup>.

Mutaciones de ganancias de función que afectan a *IL7R* se han observado en aproximadamente el 10% de los pacientes pediátricos y adultos<sup>209,274,275</sup> lo que lleva a la activación constitutiva de la señalización JAK/STAT. Además, la activación de esta vía también puede ocurrir por mutaciones en *JAK3*, presentes entre un 16-25% aproximadamente de los pacientes pediátricos y adultos<sup>209,235,275</sup>; mutaciones en *JAK1*, con frecuencias variables según la cohorte estudiada, que van desde un 1-4%<sup>209,235,275</sup> hasta un 20%<sup>209,276</sup>; y mutaciones en *STAT5*, presentes en un 8% de pacientes pediátricos y adultos<sup>275,277</sup>.

Por último, otros dos genes implicados en estas vías son *PTPN2* y *DNM2*, ambos reguladores negativos. En el caso de *PTPN2*, una fosfatasa que en condiciones normales se encarga de desfosforilar a JAK y STAT5 inhibiendo su acción y actuando como gen supresor de tumores<sup>278</sup>, se han detectado deleciones en el 7%<sup>278–280</sup> de los pacientes, prácticamente todos del subtipo TLX1<sup>279</sup>, lo que lleva a la activación constitutiva de la vía JAK-STAT<sup>278,280</sup>. Las mutaciones de pérdidas de función en *DNM2*, identificadas en el 11% de los pacientes pediátricos<sup>209</sup> y aproximadamente el 15% de los pacientes adultos<sup>235,281</sup>, llevan a un potenciamiento de la señalización de *IL7R* ya que se altera la endocitosis del mismo y aumenta la densidad del receptor en la superficie celular.

En total, alteraciones que afectan a componentes de la vía de señalización JAK/STAT y que llevan a su activación se dan en el 30% de los pacientes con LLA-T, siendo especialmente frecuentes en los subtipos TLX1, TLX3, HOXA y ETP-ALL<sup>189,209,274,275</sup>.

Un estudio mostró que las mutaciones en *IL7R* y *JAK1* confieren resistencia a glucocorticoides debido a una activación de la vía MEK-ERK y AKT, que lleva a un incremento de la expresión de agentes anti-apoptóticos (*MCL1* y *BCLXL*)<sup>282</sup>. En otro estudio se mostró que las mutaciones que afectan a la vía de señalización JAK/STAT eran especialmente frecuentes en pacientes recaídos/refractarios, indicando que estas alteraciones son un marcador de mal pronóstico en la LLA-T<sup>265</sup>.

- **Vía de señalización activada por genes RAS**

Las proteínas RAS, incluyendo las isoformas HRAS, NRAS y KRAS, son una familia de proteínas con actividad GTPasa que transforman señales procedentes de múltiples receptores de superficie celular, activando vías de señalización como PI3K/AKT y MAPK y regulando varios procesos celulares como la proliferación, diferenciación y apoptosis. Las mutaciones en los genes *RAS* activan constitutivamente estas proteínas con independencia de señales extracelulares, provocando una proliferación celular aberrante<sup>283</sup>.

En la LLA-T, *NRAS* es el gen más frecuentemente mutado de esta familia con una frecuencia de alrededor del 7% en pacientes pediátricos<sup>209,284,285</sup> y del 10% en pacientes adultos<sup>263,286</sup>. En el caso de *KRAS* las frecuencias son variables pero se encuentra mutado entre un 2-11% en el caso de pacientes pediátricos<sup>284,285</sup> y entre un 2-6% en el caso de los adultos<sup>263,286</sup>. Estas mutaciones son especialmente prevalentes en el subtipo ETP-ALL<sup>189,197</sup>, aunque también se han asociados con pacientes TLX1/TLX3 y HOXA positivos<sup>287</sup>.

Las mutaciones en *N/KRAS* están asociadas a una peor supervivencia y a un incremento de la incidencia de la recaída en pacientes adultos<sup>263,265</sup>, e identifican pacientes pediátricos con mal pronóstico, aunque no tienen repercusión en su supervivencia<sup>288</sup>. Pacientes con mutaciones en *N/KRAS* muestran resistencia a varios quimioterapéuticos<sup>222,282,289,290</sup>. Todo ello explica el hecho de que la frecuencia de estas mutaciones incremente en la recaída: 12% en el caso de *KRAS* y 27% en el caso de *NRAS*<sup>222</sup>. Por ello, son una diana de gran interés para el desarrollo de nuevas aproximaciones terapéuticas<sup>289,291,292</sup>.

Además de los genes *RAS*, también se han encontrado alteraciones en el gen *NF1*, regulador negativo de la señalización de RAS. *NF1* codifica para una proteína con actividad GTPasa que inhibe la señalización de *RAS* mediante la hidrólisis de su conformación activa, pasando de RAS-GTP a RAS-GDP<sup>293,294</sup>. En la LLA-T, se han observado microdeleciones que afectan a este gen en 3/103 y 2/27 pacientes pediátricos<sup>295,296</sup>. La repercusión funcional de estas deleciones es similar a las mutaciones activadoras en *RAS*, una activación de esta vía.

- **Reguladores del ciclo celular**

- ***CDKN2A/B***

El gen *CDKN2A*, también conocido como *INK4A* o *P16-INK4A*, y su parólogo *CDKN2B*, o *INK4B* o *P15-INK4B*, se localizan en la citobanda 9p21.3. Las proteínas codificadas por estos genes, P16 y P15, respectivamente, pertenecen a la familia de proteínas INK que se unen directamente a las ciclinas dependiente de quinasas (en inglés *Cyclin-dependent kinases*, CDK) CDK4/CDK6 inhibiendo su actividad. Por tanto, evitan que estas quinasas se unan al RB inhibiendo la progresión del ciclo celular de la fase G1 a la fase S<sup>297</sup>. Ambos genes pueden generar diferentes transcritos, entre estos destaca el transcrito de *CDKN2A* que codifica para la proteína ARF (*alternative reading frame*) que se une directamente a un regulador negativo de P53, MDM2, activando de este modo la función supresora de tumores de *TP53*, favoreciendo la apoptosis<sup>298</sup>.

En la LLA-T, se han descrito deleciones mono-alélicas y bi-alélicas de *CDKN2A/B* en un 70% de los pacientes tanto pediátricos como adultos, siendo la alteración más común de esta neoplasia<sup>299</sup>. Se considera como un evento secundario que coopera con otras mutaciones en el desarrollo de la LLA-T<sup>300</sup> y un estudio de *single cell* ha revelado que estas alteraciones serían eventos intermedios en la LLA-T<sup>301</sup>. Son especialmente frecuentes en los subtipos corticales de LLA-T (TLX1, TLX3, NKX2-1 y NKX2-2)<sup>112</sup>, mientras raramente se encuentran en el subtipo ETP-ALL<sup>189,238,302</sup>.

Además de las deleciones, el gen *CDKN2B* también puede ser inactivado por silenciamiento epigenético mediante la hipermetilación de su promotor, en este caso esta alteración es más frecuente en el subtipo EPT-ALL<sup>286</sup>. No obstante, sea cual sea el mecanismo, la inactividad de estas proteínas lleva a una proliferación aberrante de las células favoreciendo el desarrollo tumoral<sup>303</sup>.

Debido a su frecuencia, son muchos los grupos que han estudiado el impacto pronóstico de las deleciones en *CDKN2A/B*. No obstante, los resultados son heterogéneos<sup>186,281,302,304–307</sup>. La mayoría de los trabajos en cohortes pediátricas han mostrado que estas alteraciones no tienen impacto pronóstico, con la excepción del trabajo del grupo NOPHO, donde se asociaron con una baja supervivencia global<sup>308</sup>. En el caso de los pacientes adultos, la mayoría de los trabajos muestran que estas alteraciones confieren mejor pronóstico, con la excepción de la cohorte UKALL, donde no tuvieron impacto en la supervivencia global<sup>309</sup>.

#### ○ ***CDKN1B***

Este gen, también denominado p27, pertenece a la familia de inhibidores de quinasas Cip/Kip. En condiciones normales, se une a las ciclinas CDK2/CDK4/CDK6 inhibiendo el ciclo celular en la fase de transición G1-S<sup>310</sup>.

Aproximadamente un 12% de los pacientes pediátricos con LLA-T presentan deleciones en *CDKN1B*<sup>311</sup>. La inactivación de este gen conduce a una proliferación alterada de las células<sup>312</sup>. Los pacientes con deleciones en *CDKN1B* presentan un perfil de expresión HOXA+ y una desregulación de *MEF2C*, la cual se ha asociado con una peor respuesta a esteroides<sup>161</sup>.

#### ○ ***TP53***

*TP53* es uno de los principales genes supresores de tumores. En condiciones normales este FT está inactivo debido a su rápida degradación por MDM2. En condiciones de estrés celular, la actividad de MDM2 es detenida y lleva a la acumulación de P53, que induce la apoptosis y un bloqueo irreversible del ciclo celular. Alrededor del 50% de los cánceres humanos presentan inactividad de P53<sup>313</sup>.

No obstante, la frecuencia de las alteraciones de este gen, bien mutaciones o deleciones, son raras en el momento del diagnóstico en la LLA-T<sup>281,314</sup>. En un trabajo publicado recientemente, por el grupo cooperativo francés (GRAALL), se han identificado alteraciones en *TP53* en un 4% de 476 pacientes adultos y pediátricos, asociadas a un inmunofenotipo

inmaduro, un aclaramiento más lento de la ERM al final de inducción, y una mayor incidencia de recaída y menor supervivencia global, aunque no independientes de la ERM<sup>315</sup>.

La frecuencia de las alteraciones en *TP53* aumenta hasta un 20% en la LLA-T en recaída<sup>288,316,317</sup> (ver apartado 3.8.2.1).

- **RB1**

Es otro de los principales supresores tumorales. Se encarga de inhibir la actividad de FTs requeridos para la progresión del ciclo celular, en concreto produce un bloqueo en la fase G1. Su inactivación por parte de la fosforilación mediada por CDK4/6 lleva a una progresión del ciclo y división celular<sup>318</sup>.

Las deleciones en este gen, han sido identificadas en pacientes con LLA-T con frecuencias variables: 24,4% en una cohorte pediátrica<sup>319</sup>, 9,5% en una cohorte pediátrica y adultos jóvenes<sup>209</sup> y un 17% en una cohorte de pacientes adultos<sup>281</sup>.

- **Señalización de quinasas**

- **Fusión *NUP214::ABL1***

En el trabajo publicado por Graux *et al.* se identificó la fusión entre *ABL1* y *NUP214*, en forma de estructura episomal, como un nuevo mecanismo de activación de la tirosina quinasa *ABL1*<sup>320</sup>. En este mismo estudio se identificó este reordenamiento en un 6% de pacientes pediátricos<sup>320</sup> y posteriormente se ha identificado en un 4% de pacientes adultos<sup>321</sup>.

La activación constitutiva de esta tirosina quinasa lleva a la activación de *STAT5* y la vía de señalización RAS-MAPK, pero su actividad es mucho más débil que la desencadenada por la fusión *BCR::ABL1*<sup>322</sup>. De hecho, sus propiedades oncogénicas dependen de la interacción con diferentes factores y de la actividad de *LCK*, un miembro de la familia de las quinasas Src<sup>323</sup>, y por sí sola no induce el desarrollo de la leucemia<sup>324</sup>. No obstante, su co-expresión con *TLX1* lleva a un desarrollo de la LLA-T más agresiva que cuando se presenta solo una de las dos alteraciones<sup>324</sup>. De hecho, la fusión *NUP214::ABL1* es concomitante con la expresión de *TLX1* y *TLX3*<sup>209,325</sup>.

Las LLA-T con la presencia de *NUP214::ABL1* son sensibles a inhibidores de tirosina quinasa (en inglés *tyrosin kinase inhibitors*, TKI) como imatinib y dasatinib<sup>326</sup>.

- **FLT3**

*FLT3* es un gen que codifica para un receptor tirosina quinasa expresado en células hematopoyéticas inmaduras. Las alteraciones más características de *FLT3* son pequeñas mutaciones en tándem (en inglés, *internal tandem duplications*, ITD) de los aminoácidos que se encuentran dentro del dominio yuxtamembrana del receptor, descritas principalmente en LMA<sup>327,328</sup>. Esto lleva a la activación constitutiva de la quinasa *FLT3* y consecuentemente a la activación de múltiples vías de señalización activadas por la misma (Ras/MAPK, MEK/ERK y PIK3/Akt).

En el caso de la LLA-T, las alteraciones en *FLT3* ocurren de manera exclusiva en el subtipo ETP-ALL, afectando a un 14% de estos pacientes<sup>189</sup>. De hecho, la presencia de estas alteraciones define un nuevo subtipo de ETP-ALL con unas características concretas: inmunofenotipo característico (CD2+/CD5-/CD13+/CD33-), patrón de expresión génica diferente al resto de ETP-ALL (expresión aberrante de *IGFBP7*, mutaciones en *WT1*, baja expresión de *GATA3*), ausencia de mutaciones en *NOTCH1* y baja frecuencia de reordenamientos del *TCR*<sup>87,238</sup>. Además, tal y como se ha mencionado en esta introducción, un 30% de las LLA-T inmaduras con activación de BCL11B presentan alteraciones en *FLT3*<sup>85,96</sup>.

- **Reguladores epigenéticos**

- ***PHF6***

Este gen remodelador de la cromatina está inactivado por deleciones y mutaciones de pérdida de función que afectan al 15% de los casos pediátricos de LLA-T y a un 35-40% de los casos adultos<sup>329</sup>. Aparecen prácticamente solo en pacientes del sexo masculino, debido a su localización en el cromosoma X, y son especialmente frecuentes en los subtipos TLX1 y TLX3<sup>329</sup>.

Mutaciones germinales en *PHF6* causan el síndrome Börjeson–Forssman–Lehmann (BFLS)<sup>330</sup> y se ha sugerido una relación de este síndrome con una predisposición a la LLA-T<sup>331</sup>.

Aunque su función en el desarrollo de la LLA-T se desconoce, son varios los estudios que demuestran que su deficiencia favorece el inicio y desarrollo de la misma: la pérdida de *PHF6* lleva a una aceleración de la leucemia desencadenada por la expresión de *TLX3*<sup>332</sup> y disminuye el umbral necesario para desarrollar LLA-T inducidas por *NOTCH1*<sup>333</sup>. No obstante, estos estudios indican que solo la pérdida de *PHF6* no es suficiente para inducir la transformación<sup>333,334</sup>. De hecho, las mutaciones en este gen están asociadas con alteraciones en los miembros de la vía JAK/STAT (*IL7R*, *JAK1*, *JAK3* y *STAT5*) y se sabe que coopera con *JAK3* en el desarrollo de la LLA-T mediante la inhibición de la vía de señalización BAI1-MDM2-P53<sup>335</sup>.

- ***Polycomb complex (PRC2): EZH2, EED, SUZ12***

El complejo represivo Polycomb 2 (PRC2) es un regulador epigenético involucrado en la represión transcripcional génica a través de la metilación de la histona H3K27me3<sup>336</sup>. En la LLA-T, aproximadamente un 25% de los pacientes, tanto pediátricos como adultos, presentan mutaciones y deleciones que llevan a la pérdida de función de los genes *EZH2*, *EED* y *SUZ12*, codificantes de los principales componentes del complejo PRC2<sup>189,337</sup>. Estas mutaciones son más comunes en el subtipo ETP-ALL<sup>189</sup>.

Las mutaciones inactivadoras del PRC2 se han asociado a una peor respuesta al tratamiento, pues confieren a las células resistencia a la apoptosis, aunque esto no se ha asociado con una tasa de supervivencia menor<sup>338</sup>.

- **KDM6A**

*KDM6A (UTX)* es una histona demetilasa de H3K27me3 involucrada en la regulación epigenética a través de la represión de la actividad de PRC2 y tiene función de gen supresor de tumores. No obstante, también se ha descrito un posible papel oncogénico dentro del subtipo TAL1<sup>339</sup>. Mutaciones y deleciones en este gen se han descrito en el 5-15% de los pacientes con LLA-T que llevan a una pérdida de su función<sup>340-342</sup>. Debido a su localización en el cromosoma X, la inactivación de este gen contribuye al desarrollo tumoral en hombres<sup>340</sup>.

- **DNMT3A**

Las ADN metiltransferasas *DNMT1*, *DNMT3A*, y *DNMT3B* son genes que codifican a proteínas que se encargan de establecer el patrón de metilación del ADN, mediante la adición de grupos metilo a las citosinas de las islas CpG<sup>343</sup>. Se han descrito mutaciones en *DNMT3A* en un 4-18% de los pacientes con LLA-T<sup>235,344</sup>, siendo muy poco frecuentes en pacientes pediátricos (1,4%)<sup>345</sup>. De hecho, las mutaciones en *DNMT3A* están asociadas a pacientes con edades avanzadas y leucemias inmaduras, siendo especialmente frecuentes en el subtipo ETP-ALL<sup>86</sup>.

Los pacientes con mutaciones en *DNMT3A* presentan bajas tasas de remisión y un incremento en la incidencia de recaída<sup>70</sup>.

Como se ha comentado anteriormente, se han identificado mutaciones en este gen en la fracción no leucémica de pacientes adultos con LLA-T, sugiriendo una asociación con eventos de tipo CHIP<sup>70</sup>.

- **Otras alteraciones epigenéticas: TET1, TET2, IDH1, IDH2**

Las enzimas TET promueven la demetilación del ADN. La presencia de mutaciones en los genes que codifican a estas enzimas lleva a un desbalance de la homeostasis epigenética<sup>346</sup>. Mutaciones en el gen *TET1* se han reportado en un 6-14% de pacientes pediátricos y adultos con LLA-T<sup>347,348</sup>. La interacción entre *TET2* inactivado y *DNMT3A* mutado altera el perfil de metilación del ADN en un modelo de ratón, alterando la función de varios supresores tumorales<sup>349</sup>.

Las mutaciones en los genes *IDH1* e *IDH2* llevan a la producción de un onco-metabolito que inhibe la actividad enzimática de TET, llevando también a una desregulación del equilibrio epigenético<sup>350</sup>. Mutaciones en *IDH1* e *IDH2* han sido identificadas en el 2-10% de pacientes adultos con un inmunofenotipo inmaduro, subgrupo asociado a mal pronóstico<sup>197,281,351</sup>.

- **Proteínas ribosomales y con función en la traducción**

La transformación y proliferación de las células tumorales requiere de un aumento de la síntesis de proteínas<sup>352</sup>. En pacientes con LLA-T se han identificado alteraciones en *CNOT3*, *RPL5*, *RPL10* y *RPL22*, genes que codifican para proteínas ribosomales. En concreto *RPL5*, *RPL10* y *RPL22* codifican para componentes de la subunidad 60S.

Mutaciones en el gen *RPL5* se han reportado en el 2% de pacientes tanto pediátricos como adultos con LLA-T<sup>347</sup>.

El gen *RPL10* se encuentra alterado en el 8,2% de pacientes pediátricos y en el 1% de adultos<sup>347</sup>. La mutación R98S es la más frecuente e impide el correcto ensamblaje del ribosoma afectando a la proliferación celular<sup>353</sup>. En concreto en la LLA-T, se ha visto que las células leucémicas con esta mutación tienen una sobre-expresión de *BCL2* lo que previene la apoptosis<sup>354</sup>. Además, las mutaciones en *RPL10* *up-regulan* componentes de la vía de señalización JAK-STAT llevando a una hiper activación de esta vía y a una estimulación de citoquinas, favoreciendo la progresión de la LLA-T.<sup>355</sup>

En el caso de *RPL22* se han reportado alteraciones que llevan a la inactivación del gen en un 10% de los pacientes pediátricos, siendo principalmente deleciones. Esta inactivación predispone a los progenitores de linaje T a sufrir una transformación oncogénica<sup>356</sup>.

En el gen *CNOT3*, se han identificado mutaciones en menos del 1% de pacientes pediátricos y en el 8% de pacientes adultos. Las mutaciones en este gen sugieren que actúa como supresor tumoral. De hecho, estudios en un modelo de *Drosophila* mostraron que la reducción de la expresión de Not3 era suficiente para producir una transformación de las células, confirmando su potencial oncogénico<sup>347</sup>.

Las alteraciones que afectan a estos genes no se han asociado con ningún subtipo molecular de LLA-T.

- **Otras alteraciones citogenéticas**

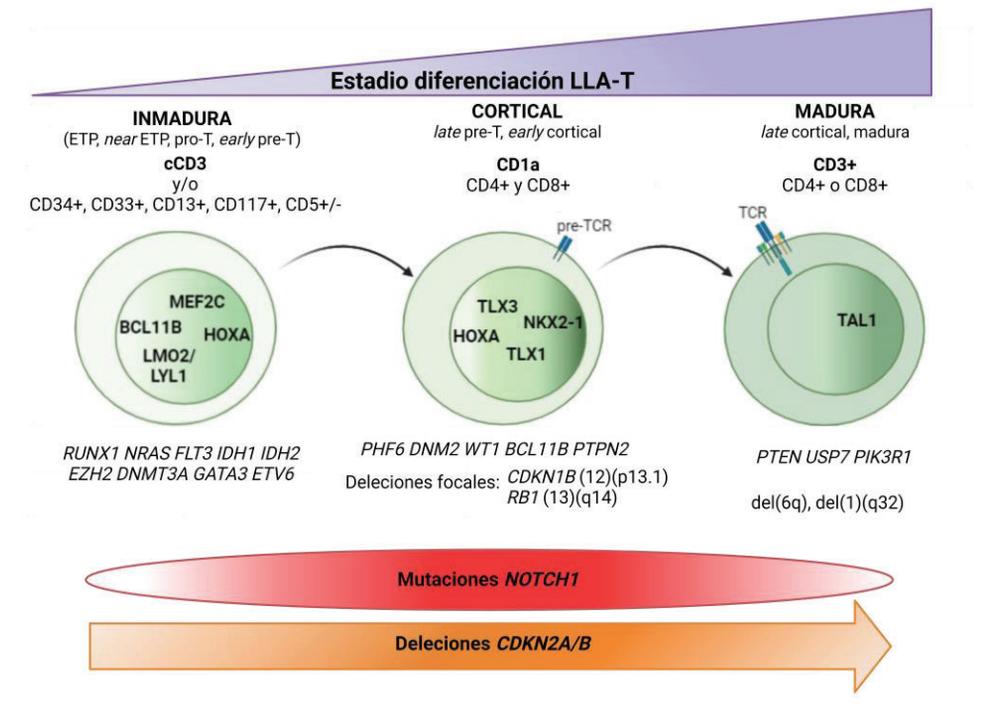
- ***SET::NUP214***

La fusión *SET::NUP124*, consecuencia de la deleción del(9)(q34.11q34.13), fue descrita por Van Vlierberghe *at al.* en 3 pacientes pediátricos de LLA-T<sup>357</sup>. Fue en este mismo trabajo donde se describió que el gen de fusión resultante se une a regiones promotoras de genes *HOXA* llevando a su activación transcripcional y, por tanto, contribuyendo en la patogénesis de la LLA-T favoreciendo el bloqueo de la diferenciación<sup>357</sup>.

- ***Del(6q)***

Esta alteración citogenética se ha identificado en un 10-20% de pacientes pediátricos<sup>358,359</sup> y 10% en pacientes adultos<sup>309</sup>. Fue en el trabajo publicado por Gachet *et al.*<sup>360</sup> donde identificaron que estas deleciones llevaban a la infra-regulación de dos genes (*SNHG5* y *SYNCRIP*) implicados en el procesado del ARN mensajero y la biogénesis ribosomal. Son comunes en el subtipo TAL1<sup>360</sup> y se han asociado a mala respuesta al tratamiento<sup>311</sup>.

Así pues, como se ha ido mostrando a lo largo de este apartado, las diferentes alteraciones secundarias características de la LLA-T (alteraciones tipo B) se encuentran asociadas a un subtipo concreto de LLA-T definido por el perfil de expresión resultante de la activación de FT (alteraciones tipo A) que, a su vez, está relacionado con el inmunofenotipo que presentan los blastos. Por lo tanto, inmunofenotipo, perfil de expresión y alteraciones secundarias están asociados y ayudan a definir los diferentes subtipos de LLA-T como se muestra en la figura 5:



**Figura 5. Alteraciones genéticas que definen los subtipos de LLA-T.** Representación esquemática de los tres principales subtipos de LLA-T de acuerdo con el estadio de maduración donde se produce el bloqueo de la diferenciación. Los marcadores inmunofenotípicos definitorios de cada subtipo se marcan en negrita en la parte superior, junto a otros marcadores propios de cada subtipo. El proceso de maduración del TCR también se representa. La presencia del TCR en la superficie celular del blasto también es un marcador de subtipo maduro. Los factores de transcripción activos definitorios de cada subtipo se representan en el núcleo siguiendo el orden en el proceso de maduración. Los genes frecuentemente mutados en cada subtipo seguidos de las alteraciones en el número de copias se muestran en la parte de abajo. La distribución de las mutaciones en *NOTCH1* y deleciones en *CDKN2A/B* se indican en la parte inferior de la figura. Adaptada de Genescà y González-Gil, 2022<sup>361</sup>.

### 3.6. FACTORES PRONÓSTICOS EN LA LLA-T

La identificación de factores pronósticos que permitan evaluar el riesgo de cada paciente es fundamental a la hora de diseñar protocolos de tratamiento como también poder estratificar a los pacientes. En el último protocolo de tratamiento pediátrico europeo (ALL TOGETHER, NCT04307576) los pacientes se estratifican en bajo riesgo, riesgo intermedio bajo, riesgo intermedio alto, alto riesgo quimioterapia y alto riesgo TPH (trasplante de progenitores hematopoyéticos), en función de su evolución al tratamiento de inicio.

En el caso de los pacientes adultos, el último protocolo de tratamiento de LLA del grupo PETHEMA (LAL19), la estratificación también depende de la evolución de la enfermedad cuantificada por la de ERM tras el primer bloque de inducción y de la genética en el momento del diagnóstico.

En el caso concreto de la LLA-T, el hecho que los pacientes se diagnostiquen con esta leucemia se ha considerado, tradicionalmente, un factor de alto riesgo en protocolos pediátricos. No obstante, el estudio y la propuesta de nuevos factores pronósticos es la base para ir hacia una terapia personalizada donde se tengan en cuenta las características basales de cada paciente como también las características genéticas y fenotípicas de la leucemia.

De este modo, los factores pronósticos se dividen en: factores derivados del paciente, factores derivados de la enfermedad y factores pronósticos de respuesta al tratamiento. No obstante, estos factores son los propuestos para la LLA en general y salvo para el inmunofenotipo no se han propuesto factores pronósticos específicos para la LLA-T.

### 3.6.1 Factores pronósticos del paciente adulto

- **Edad**

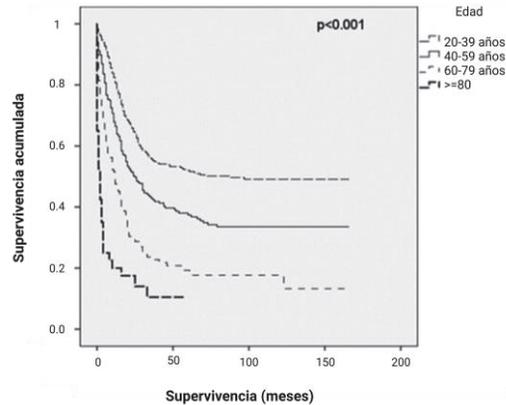
La edad es uno de los parámetros clínicos más relevantes en la estratificación del riesgo de los pacientes con LLA. Históricamente, la edad ha permitido establecer dos grandes grupos con marcadas diferencias en las tasas de supervivencia: los pacientes pediátricos (<18 años) alcanzan tasas de supervivencia superiores al 90%, mientras que los adultos difícilmente superan el 40%<sup>362</sup>.

Además, existe un grupo intermedio compuesto por adolescentes y adultos jóvenes (en inglés, *adolescents and young adults*, AYA) (15-39 años), que presentan un pronóstico intermedio, mostrando tasas de supervivencia de alrededor del 60-70%<sup>362</sup>.

Estudiando específicamente el subtipo LLA-T, se confirma lo anteriormente citado ya que existen diferencias en las tasas de supervivencia entre las cohortes pediátricas frente a cohortes adultas<sup>363,364</sup>. Además, también hay diferencias entre los distintos grupos de edad dentro de las cohortes de adultos, disminuyendo la probabilidad de supervivencia con el aumento de la edad<sup>365</sup>.

De este modo, en un estudio realizado en una cohorte de 1141 pacientes adultos (mediana de edad de 35 años) en EE. UU. se observaron diferencias en las supervivencias globales entre grupos de edades, siendo la probabilidad de supervivencia global a los 5 años para el grupo de <40 años de 51,9%; para el grupo de entre 40-59 años de 37,3%; para el grupo de entre 60-79 años de 19,2% y para el grupo de >80 años de 0% (Fig. 6)<sup>365</sup>.

No obstante, aunque la edad marca diferencias de supervivencia entre los pacientes no es un factor pronóstico independiente cuando se incluye la variable ERM<sup>366</sup>.



**Figura 6. Supervivencia global pacientes LLA-T en EE. UU. estratificados por edad.** Adaptada de Guru Murthy *et al.* 2019<sup>365</sup>.

### 3.6.2 Factores pronósticos de la enfermedad

- **Nº leucocitos al diagnóstico**

Es un indicador de la carga tumoral y la capacidad proliferativa de la leucemia. Es otro de los factores pronósticos clásicos. En la LLA-T se ha descrito que los pacientes con un  $WBC > 100 \times 10^9$  leucocitos/l tienen peor pronóstico<sup>309</sup>. No obstante, las ETP-ALL se caracterizan por tener un WBC inferior presentando peor supervivencia que los grupos con un recuento mayor<sup>85</sup>.

A pesar de que hay diferencias en la supervivencia de los pacientes con distintos valores de WBC, esta variable no es un factor pronóstico independiente cuando se incluye la variable ERM<sup>366</sup> y actualmente se cuestiona su valor pronóstico.

- **Inmunofenotipo**

Como se ha explicado anteriormente, los diferentes estadios de maduración característicos de cada subtipo de LLA-T también se han asociado al pronóstico (**Ver tabla 2**). A nivel clínico y de diagnóstico, el subgrupo inmunofenotípico más inmaduro de LLA-T (ETP-ALL) se ha asociado a peores tasas de supervivencia<sup>85,367</sup>.

No obstante, la mayoría de estudios realizados, tanto en cohortes pediátricas como adultas, han mostrado que el mal pronóstico de los pacientes con este inmunofenotipo puede mejorar con esquemas de tratamiento más intensivos<sup>368-371</sup> o con un Alo-TPH<sup>90,369,372</sup>.

En el protocolo de LLA de PETHEMA actual (LAL19) los pacientes con este subtipo realizan un protocolo de tratamiento diferente al resto de pacientes siguiendo un esquema tipo FLAG-IDA inspirado de los tratamientos de la LMA.

- **Citogenética**

En la LLA-T, como se ha explicado anteriormente, no hay una clasificación citogenética establecida. No obstante, el trabajo publicado por Moorman *et al.* mostró que los pacientes

con LLA con un cariotipo complejo, definido por la presencia de más de 5 alteraciones, tenían peores tasas de supervivencia<sup>373</sup>. En otro estudio publicado más recientemente, donde se analizaron 216 pacientes adultos con LLA-T, se mostró que los pacientes con 3 o más alteraciones en el cariotipo tenían peor pronóstico con independencia del valor de la ERM<sup>374</sup>. Sin embargo, el cariotipo no es un marcador que se utilice actualmente en el manejo clínico de los pacientes.

- **Genética**

Como se ha visto a lo largo de los apartados anteriores el papel que tienen las alteraciones genéticas en el inicio y desarrollo de la LLA-T es indudable. Por ello, es posible pensar que estas alteraciones pueden ser marcadores de pronóstico. No obstante, dada la heterogeneidad genética de la LLA-T, hasta ahora ha sido difícil lograr un consenso entre los diferentes grupos de trabajo.

Debido a que las mutaciones que afectan a los genes *NOTCH1/FBXW7* son las más frecuentes en los pacientes con LLA-T, su impacto pronóstico ha sido evaluado en muchos trabajos. Algunos las han asociado con una mejor respuesta al tratamiento y mejor supervivencia<sup>237,375</sup>. Sin embargo, no todos los estudios muestran resultados estadísticamente significativos y hay una heterogeneidad entre las tasas de supervivencia cuando se comparan protocolos de tratamiento similares. El trabajo de Trinquand *et al*<sup>263</sup> fue pionero en proponer una estratificación oncogenética del riesgo de los pacientes con LLA-T indicando que los pacientes con mutaciones en *NOTCH1/FBXW7* sin alteraciones en los genes *PTEN* o *NKRAS* representaban un grupo de buen pronóstico con tasas de supervivencia mayores y tasas de recaída inferiores.

- **Infiltración del SNC**

La presencia de blastos en el líquido cefalorraquídeo se da en aproximadamente en un 10% de los pacientes adultos con LLA-T, un porcentaje mayor que en el caso de la LLA-B<sup>74</sup>. Una infiltración de los blastos en el SNC está asociada a tasas de supervivencia menores y mayor probabilidad de sufrir una recaída y se ha considerado un factor de mal pronóstico<sup>74,376,377</sup>. Sin embargo, con la adecuada profilaxis del SNC con terapia intratecal y terapia sistémica que atraviesa la barrera hematoencefálica, el valor pronóstico de la infiltración del SNC ha desaparecido.

### 3.6.3 Factores pronósticos de respuesta al tratamiento

- **ERM**

Todos los factores pronósticos comentados anteriormente se evalúan en el momento del diagnóstico. No obstante, una vez iniciado el tratamiento se debe reevaluar el riesgo y la respuesta del paciente, que guiará la intensificación de la terapia. En este sentido la ERM es el factor que aporta más información. Este parámetro muestra la cantidad de células leucémicas ocultas presentes en el paciente durante el tratamiento y se ha convertido en el factor pronóstico más potente de la LLA, y en la mayoría de los casos es el único factor pronóstico<sup>378-380</sup>.

En la actualidad, los protocolos de tratamiento de LLA están guiados por la respuesta de la ERM<sup>76,380,381</sup> (Tabla 3). El valor de  $10^{-4}$  (0,01%) o inferior es el utilizado a nivel internacional para definir la remisión molecular<sup>76</sup>. De este modo, los pacientes con un valor de ERM > 0,01% tienen una ERM positiva mientras que los pacientes con ERM < 0,01% tienen una ERM negativa. Los pacientes con ERM positiva al final del tratamiento de inducción se han asociado con un alto riesgo a recaer y se consideran candidatos a recibir TPH alogénico tras la consolidación. Por el contrario, la indicación del TPH en los pacientes con ERM negativa tras la inducción es más cuestionable y en estudios recientes se consideran candidatos a recibir quimioterapia de consolidación y mantenimiento<sup>380,382</sup>. No obstante, la cinética de aclaramiento de la ERM se considera un parámetro más importante que la simple determinación de la ERM tras la inducción. En los pacientes con LLA-T parece ser más lenta que en el caso de la LLA-B, y se ha constatado en pacientes pediátricos que, aunque el valor negativo de la ERM al final de inducción es un factor favorable, pacientes con ERM positiva inicial que alcanzan una ERM negativa tras consolidación también tienen buen pronóstico<sup>383</sup>.

**Tabla 3. Definiciones de respuesta al tratamiento en función del valor de la ERM**

Terminología	Definición
Respuesta Completa Hematológica (RC Hem)	<ul style="list-style-type: none"> <li>Células leucémicas no detectables en MO, SP, LCR (&lt; 5% blastos) por microscopía</li> </ul>
Respuesta Completa Molecular (ERM <sup>-</sup> ) (RC mol)	<ul style="list-style-type: none"> <li>Paciente en RC Hem</li> <li>ERM no detectada por qPCR (sensibilidad <math>\geq 10^{-4}</math>)</li> </ul>
Respuesta Molecular (menos que RC mol)	<ul style="list-style-type: none"> <li>Paciente en RC Hem, no en RC mol</li> <li>ERM no cuantificable por qPCR (<math>&lt;10^{-4}/0,01\%</math>)</li> <li>Detectable mediante CFM 8-12 colores</li> </ul>
Fallo Molecular / ERM <sup>+</sup>	<ul style="list-style-type: none"> <li>Paciente en RC Hem, no en RC mol</li> <li>ERM detectada por qPCR (<math>&lt;10^{-4}/0,01\%</math>)</li> <li>Detectable por CFM</li> </ul>
Recaída Molecular	<ul style="list-style-type: none"> <li>Paciente todavía en RC Hem, habiendo alcanzado previamente RC Mol</li> <li>Pérdida de RC Mol (<math>\geq 10^{-4}/0,01\%</math>)</li> <li>Detectable por CFM</li> </ul>
Recaída	<ul style="list-style-type: none"> <li>Pérdida de RC Hem</li> <li>Recaída Hematológica (&gt; 5% blastos en MO)</li> <li>Recaída extramedular (SNC, otro tejido)</li> </ul>

MO: Médula ósea; SP: Sangre periférica; LCR: líquido cefalorraquídeo, ERM: Enfermedad residual medible; CFM: citometría de flujo multiparamétrica; SNC: sistema nervioso central. Adaptada de Hoelzer *et al.* 2016<sup>76</sup>

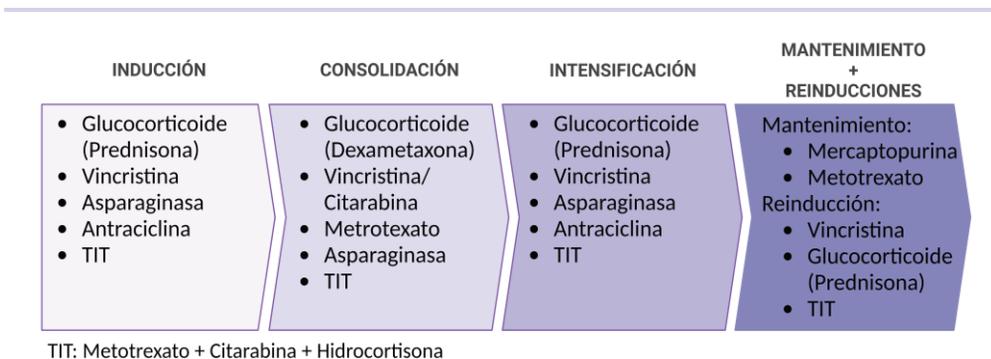
La monitorización de la ERM se puede realizar por citometría de flujo, por PCR cuantitativa o por NGS. En el primer caso, se analizan el número de células con un inmunofenotipo aberrante, identificado en el diagnóstico, en las muestras de seguimiento. La sensibilidad de la técnica depende del número de eventos que se analizan pudiendo llegar a una detección de hasta  $10^{-5}$  -  $10^{-6}$ <sup>384,385</sup>. En la PCR cuantitativa, en el caso de la LLA-T, se cuantifican los reordenamientos de los genes del receptor de células T (*TCRG*, *TCRD*, *TCRB*, *TCRA*)<sup>386</sup>. La sensibilidad de la PCR cuantitativa es de hasta  $10^{-5}$ <sup>387</sup>. Finalmente, la NGS es otra técnica

que actualmente también se utiliza para evaluar la ERM. Las alteraciones a identificar serían las mismas que las evaluadas por PCR cuantitativa, es decir, en el caso de la LLA-T los reordenamientos del TCR, principalmente. La sensibilidad de la NGS puede llegar hasta el orden de  $10^{-6}$  o  $10^{-7}$ <sup>388</sup>.

### 3.7. TRATAMIENTO

Son múltiples los estudios que demuestran que los protocolos de tratamiento de tipo-pediátrico y basados en el uso de quimioterapia intensiva mejoran el pronóstico de los pacientes adultos tanto de LLA-B como LLA-T<sup>389-393</sup>. Es por ello, que actualmente los protocolos de tratamiento de LLA adultos, para ciertas edades, se basan en estos esquemas pediátricos o de inspiración pediátrica. En España, los adultos siguen protocolos diseñados por el grupo PETHEMA (Programa Español de Tratamientos en Hematología) que asignan el tipo de tratamiento post-consolidación (quimioterapia o TPH) en función de la dinámica de aclaramiento de la ERM y, recientemente, del riesgo genético. Estos protocolos son comunes para todos los pacientes con LLA Ph<sup>-</sup> tanto LLA-B como LLA-T.

Actualmente, los pacientes con LLA en España siguen el protocolo LAL19. Los pacientes incluidos en esta tesis han sido tratados según los protocolos Alto Riesgo 2003 (AR03) [NCT00853008] y Alto Riesgo 2011 (AR11) [NCT01540812], que presentan una alta coincidencia en su diseño general. En la figura 7 se muestra un esquema de las principales fases tratamientos y los agentes quimioterapéuticos utilizados en el tratamiento de la LLA.



**Figura 7. Principales fases del tratamiento y agentes quimioterapéuticos utilizados en el tratamiento de la LLA.** Adaptada de Malard y Mohty, 2020<sup>394</sup>

- **Prefase**

Es una fase previa al inicio del tratamiento basada en el uso de glucocorticoides como la prednisona. El objetivo de este bloque es reducir la carga leucémica y dar tiempo a completar la caracterización genética y molecular al diagnóstico de la LLA para administrar un tratamiento más específico. En esta fase suele administrarse una dosis de tratamiento intratecal (TIT).

- **Inducción**

El objetivo de la inducción es reducir al máximo la cantidad de células leucémicas y restaurar la hematopoyesis normal, adquiriendo la remisión completa, es decir, presentar menos de un 5% de blastos leucémicos en la médula ósea. Tiene una duración aproximada de 4 a 5 semanas.

La terapia de inducción generalmente incluye glucocorticoides (prednisona, prednisolona o dexametasona), vincristina (alcaloide citotóxico inhibidor de la mitosis), asparaginasa (enzima citotóxica) y una antraciclina (antibiótico antitumoral) que generalmente suele ser la daunorrubicina. Tras esta fase se evalúa la respuesta morfológica y la ERM por citometría de flujo. Los pacientes con ERM negativa al final de la inducción generalmente presentan mayor probabilidad de supervivencia<sup>395</sup>.

- **Segunda Inducción**

Los pacientes que no alcanzan la remisión completa o una ERM negativa tras el primer bloque de inducción generalmente realizan un tratamiento de rescate para intentar alcanzar la remisión.

Este bloque consiste en la combinación de las drogas fludarabina (análogo de la purina), altas dosis de la citarabina Ara-C, la antraciclina idarubicina y G-CSF (factor estimulante de colonias de granulocitos), conocido como FLAG-IDA. En el caso de los pacientes con el subtipo ETP-ALL en el último protocolo de PETHEMA (LAL19), se propone que inicien el tratamiento con un esquema tipo FLAG-IDA en lugar de con el bloque de inducción convencional.

- **Consolidación**

La consolidación se inicia inmediatamente después de alcanzar la remisión completa y tiene como objetivo eliminar la enfermedad residual. La intensidad y duración del tratamiento depende de la respuesta del paciente tras el primer bloque de inducción: los pacientes que han aclarado bien tras este primer bloque reciben 3 bloques de consolidación mientras que los que no han aclarado bien o no alcanzan la remisión completa reciben 1 bloque de consolidación, tras el que se realiza un TPH.

Estos bloques incluyen la administración de glucocorticoides (dexametasona generalmente), vincristina, asparaginasa y altas dosis del anti-metabolito metotrexato que se intercala con dosis intermedias de Ara-C. A continuación, se puede administrar un tratamiento similar al de la inducción llamado reinducción, seguido de consolidación tardía, si el paciente mantiene un buen aclaramiento de ERM. El tratamiento de consolidación tiene una duración de 3 a 6 meses.

- **Quimioterapia de Intensificación**

Después de la consolidación, se incluye una fase de intensificación con el objetivo de eliminar cualquier resto de células leucémicas. En este bloque se combina metotrexato y Ara-C junto a citostáticos usados en la inducción. Tras las fases de consolidación e intensificación hay una nueva evaluación de la ERM y de su valor dependerá los pasos siguientes del protocolo.

- **Mantenimiento**

Este bloque tiene una duración de entre 18 a 24 meses. El principal quimioterapéutico es la mercaptopurina (análogo del anti-metabolito purina) que se administra en dosis diarias y se combina con dosis semanales de metotrexato. También se aplican ciclos de reinducción que incluyen prednisona y vincristina. Durante la semana de administración del ciclo de reinducción se suspende la quimioterapia de mantenimiento.

- **Tratamiento de profilaxis de la leucemia en el sistema nervioso central (SNC)**

Junto a todos los bloques anteriormente citados se recomienda el tratamiento de profilaxis para el SNC. Tradicionalmente, la irradiación craneal era el método estándar. Sin embargo, su asociación con efectos adversos ha favorecido su desaparición<sup>396</sup>. La estrategia actual incluye quimioterapia intensiva triple intratecal combinándose metotrexato, Ara-C e hidrocortisona.

- **Trasplante de progenitores hematopoyéticos**

El uso del TPH en los protocolos de tratamiento de la LLA supone un gran debate pues es un tratamiento eficaz, pero también es el que mayor efectos adversos tiene para el paciente<sup>397</sup>. En la práctica clínica actual, la decisión de realizar un trasplante depende del valor de la ERM al final de la inducción y la consolidación. La tendencia es a trasplantar cada vez menos y hacerlo únicamente en aquellos pacientes con ERM positiva o que presenten características genéticas de alto riesgo.

### 3.8. EVOLUCIÓN CLONAL Y RECAÍDA

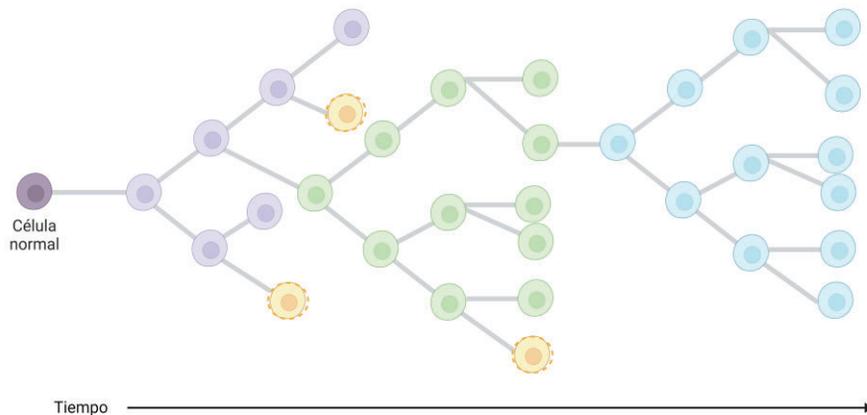
La optimización de los protocolos terapéuticos ha conseguido mejorar las tasas de supervivencia y reducir la mortalidad relacionada con el tratamiento de manera significativa en pacientes de todas las edades. Sin embargo, aproximadamente un 20% de pacientes pediátricos<sup>383</sup> y un 40% de pacientes adultos<sup>304,364</sup> sufrirán una recaída. La supervivencia global de los pacientes adultos tras la recaída disminuye hasta 20-30%<sup>364,398</sup>. Por tanto, la intensificación de la quimioterapia no es suficiente y es necesario buscar marcadores predictivos de recaída y dianas terapéuticas que mejoren estos datos.

Para estudiar las posibles causas de la recaída es necesario partir del origen de la leucemia y conocer cómo evoluciona a lo largo del tiempo. Fue el grupo de J. E. Dick quien identificó por primera vez la población celular denominada células iniciadoras de la leucemia (en inglés *leukemic initiating cells*, LIC), una población CD34<sup>+</sup>CD38<sup>-</sup> aislada de pacientes con LMA con capacidad de iniciar la enfermedad al trasplantarse en ratones inmunodeprimidos<sup>399,400</sup>. En el caso de la LLA-T, se han realizado diversos estudios con el objetivo de caracterizar esta población LIC utilizando modelos de ratones de LLA-T genéticamente modificados (modelo deficiente de *PTEN*<sup>257</sup>, modelos SCL/LMO<sup>159,401</sup> y modelos de *NOTCH1*<sup>402-404</sup>), como también partiendo de muestras de pacientes con LLA-T<sup>405-407</sup>. Los estudios en los modelos de ratón mostraron que solo una pequeña fracción de células presentes en la leucemia primaria tenía la capacidad de propagar/iniciar la enfermedad después de pases en ratones, coincidiendo con lo observado en la LMA<sup>399,400</sup>. No obstante, los intentos de aislar estas poblaciones celulares y caracterizarlas a nivel de marcadores de superficie fueron fallidos, y para cada

uno de los modelos, la población identificada como LIC fue diferente, mostrando la heterogeneidad que presenta la LLA-T<sup>159,257,401-404</sup>. Lo mismo ocurrió para las células leucémicas aisladas de muestras primarias, en este caso habiendo también diferencias en función de los diferentes grupos de edad<sup>405-407</sup>. Debido a estas dificultades, también halladas en otros tipos de leucemia, fue también el grupo de J.E. Dick quien dictó las premisas que tenía que tener una población celular para ser considerada LIC: (i) células capaces de generar la leucemia en ratones xenoinjertados derivados de pacientes (en inglés, *patient derived xenograft*, PDX) (ii) capacidad de autorrenovación tras diferentes pases en PDX, y (iii) capacidad de generar células hijas que pueden poseer capacidad proliferativa pero que son incapaces de establecer o mantener el clon leucémico tras diferentes pases<sup>408</sup>.

### 3.7.1 Evolución clonal

Los estudios citogenéticos realizados entre las décadas de 1960 y 1970 fueron los primeros en mostrar la diversidad cromosómica y los cambios que se producían a lo largo del tiempo en la leucemia<sup>409</sup>. Estos llevaron a P.C. Nowell a plantear la teoría de que el cáncer, aunque se originaría a partir de una única célula, generaría una progenie genéticamente diversa que se sometería a sucesivos procesos secuenciales de selección resultando en la expansión de los clones con mayor capacidad proliferativa, mientras que aquellos con menor capacidad proliferativa desaparecerían. El modelo de evolución clonal de P.C. Nowell sugiere una evolución de tipo lineal (Fig. 8), en el cual hay una dominancia de un solo con o subclon<sup>410</sup>.



**Figura 8. Modelo de evolución lineal propuesto por Nowell.** Esquema simplificado del modelo propuesto por Peter Nowell<sup>410</sup>. Los diferentes colores representan las mutaciones adquiridas, excepto el clon de color amarillo que presentan las células que mueren. Adaptada de Greaves, 2016<sup>411</sup>.

No obstante, en los últimos años el desarrollo de las técnicas genómicas ha permitido realizar múltiples estudios, mostrando que la evolución clonal del cáncer es un proceso más dinámico y ramificado. En el caso de la LLA, la estrategia seguida en la mayoría de estos estudios ha sido la comparación del perfil mutacional de muestras al diagnóstico y a la recaída, principalmente en muestras de pacientes pediátricos. Fue el trabajo de Mullighan *et al.*<sup>412</sup> el

pionero en realizar este tipo de aproximación. En dicho estudio se analizaron las diferencias en las alteraciones en el número de copias en muestras pareadas al diagnóstico y recaída en pacientes pediátricos, mostrándose que en el 90% de los casos el clon de la recaída no era idéntico al clon del diagnóstico. En concreto, en el 34% de los casos el clon de la recaída había adquirido nuevas alteraciones respecto al del diagnóstico sugiriendo una evolución a partir de este. En cambio, en el 52% de los casos el clon de la recaída además de haber adquirido nuevas alteraciones también había perdido otras respecto al diagnóstico, sugiriendo que el clon de la recaída había evolucionado de un clon filogenéticamente anterior al del diagnóstico.

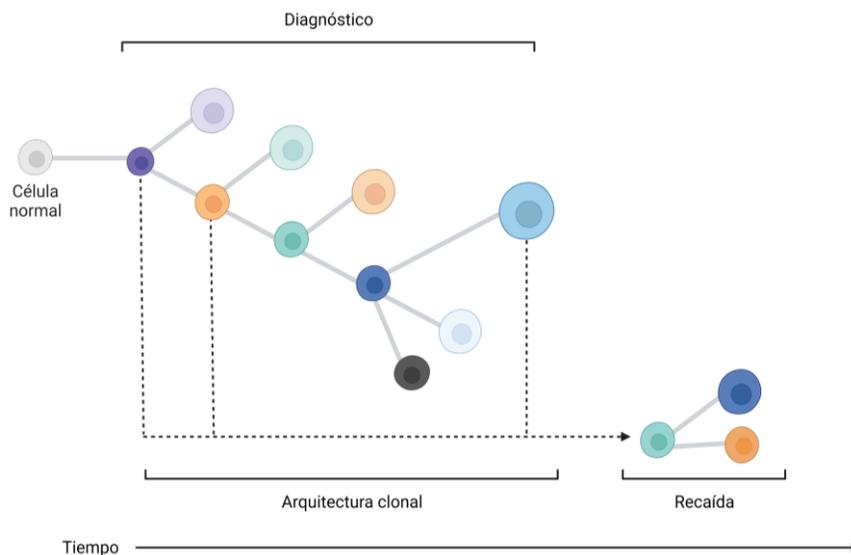
En un estudio posterior se compararon las alteraciones presentes en muestras primarias de LLA-T, principalmente de pacientes pediátricos, en el momento del diagnóstico con las presentes en las células leucémicas obtenidas tras la amplificación de estas mismas muestras en ratones PDX. Los resultados mostraron que las células derivadas del PDX presentaban alteraciones genéticas adicionales en oncogenes y genes supresores de tumores respecto a las del diagnóstico. Dado que esas células derivaban de muestras al diagnóstico, los clones amplificados en el PDX tendrían que haber derivado de los presentes en el diagnóstico. De hecho, estudiando a mayor profundidad estas alteraciones en el momento del diagnóstico se detectaron a bajos niveles, confirmándose la heterogeneidad clonal de la LLA-T. Además, comparando también las alteraciones de las mismas muestras en el momento de la recaída, se observó que el perfil mutacional de las células derivadas del PDX se asemejaba más al de la recaída, mostrando la existencia de una presión selectiva en ambos tipos de muestras que llevaría a una evolución clonal similar<sup>413</sup>.

A partir de estos trabajos, han sido varios los que se han desarrollado con el mismo fin. Todos coinciden en que la mayoría de las leucemias que aparecen en la recaída han seguido una evolución clonal ramificada a partir de un clon minoritario presente en el diagnóstico y en la mayoría de los casos de un clon ancestral al clon del diagnóstico<sup>222,414-417</sup>. En todos estos estudios se analizaron conjuntamente muestras de LLA-T y LLA-B de pacientes pediátricos, con la excepción del trabajo de Sentís *et al.*, donde se analizaron 19 tripletes de pacientes adultos con LLA-T<sup>415</sup>.

Recientemente, gracias al desarrollo de las tecnologías de *single cell* se ha podido estudiar en detalle y a nivel de célula individual la composición y evolución de los clones en la LLA-T, ya que hasta ahora se hacía mediante estimaciones partiendo de los valores de las VAF de las mutaciones del total del ADN de la muestra analizada. De este modo, mediante la secuenciación de muestras de 4 pacientes pediátricos con LLA-T en el momento del diagnóstico, se observó que todos presentaban un grupo de células dominante altamente mutado que representaba entre 28-94% de las células secuenciadas, y varios grupos más pequeños con menos mutaciones, confirmando la heterogeneidad clonal de la LLA-T<sup>301</sup>. En otro estudio utilizando esta tecnología en 5 muestras de pacientes pediátricos de LLA-T al diagnóstico, durante el tratamiento y en la recaída también se confirmó la heterogeneidad clonal en el momento del diagnóstico con la presencia de un clon mayoritario y varios clones minoritarios<sup>418</sup>. Estudiando la evolución de los clones durante el tratamiento se observó que clones minoritarios del diagnóstico evolucionaban a clones mayoritarios en etapas

posteriores donde también aparecían clones nuevos. Además, en uno de estos estudios se observó que en algunos pacientes las mutaciones presentes en el diagnóstico se identificaban en progenitores multipotentes y mieloides confirmando el origen en progenitores multilínea en algunas LLA-T<sup>301</sup>. Estos resultados coinciden con lo observado por Mel Greaves quien demostró la presencia de clones pre-leucémicos que presentaban aneuploidías y genes de fusión aberrantes durante la etapa uterina de pacientes pediátricos con LLA-B<sup>53</sup>. No obstante, estas alteraciones no serían suficientes para el desarrollo de la LLA y se requeriría de eventos secundarios, como se ha mencionado anteriormente<sup>419</sup>. En el caso de la LMA y síndromes mielodisplásicos (SMD) también se ha sugerido la presencia de clones pre-leucémicos dando lugar a procesos de CHIP<sup>420</sup>.

Este conjunto de estudios demuestra que las poblaciones LIC presentan una heterogeneidad clonal y genómica dentro del mismo paciente lo que ha debilitado la perspectiva de aislarlas según su perfil fenotípico, como se comentaba anteriormente, ganando peso el estudio de su perfil mutacional y las repercusiones que estas mutaciones pueden tener en las funciones celulares. Muchos de los clones que componen la LLA ya tendrían un origen pre-leucémico. El conjunto de clones que forman la leucemia, a diferencia de lo que proponía Nowell *et al.*<sup>410</sup>, seguirían trayectorias evolutivas complejas y ramificadas con cierto paralelismo al modelo de evolución propuesto por Darwin (Fig.9). Según este modelo, ciertos clones adquirirían mutaciones que les aportarían ventajas evolutivas para adaptarse mejor al nicho hematopoyético y seguir proliferando<sup>421</sup>. Con el tiempo la presencia de fuerzas selectivas llevaría a la selección de los clones mejor adaptados desencadenando una recaída.



**Figura 9. Modelo de evolución ramificada.** Las células de diferentes colores y tamaños representan los subclones de la población celular. Las líneas discontinuas indican que la recaída puede originarse de diferentes subclones. Adaptada de Greaves, 2016<sup>411</sup>

### 3.7.2 Recaída

Se han propuesto diferentes mecanismos como responsables de favorecer la recaída en la leucemia, como por ejemplo, la presencia de células madre inactivas, quimiorresistentes y con mayor capacidad de autorenovación<sup>422</sup>; el efecto del nicho ejerciendo una función quimio protectora<sup>423,424</sup>; y la selección de alteraciones genéticas secundarias que promueven la resistencia a la quimioterapia y que se explican en detalle a continuación.

#### 3.7.2.1 Mutaciones de recaída

Entre los genes predominantemente mutados en la recaída, destaca especialmente ***NT5C2*** que codifica a una enzima que cataliza la desfosforilación de monofostatos de nucleótidos de 6-hidroxipurina, inactivándolos. Entre estos se encuentran la 6-mercaptopurina (6-MP) y 6-thioguanina (6-TG) dos análogos de nucleósidos usados comúnmente en el tratamiento de la LLA<sup>425</sup>. Estudios en cultivos celulares han mostrado que estas mutaciones confieren resistencia a estos dos compuestos<sup>426-428</sup>. El análisis de muestras en recaída ha demostrado la presencia de mutaciones de ganancia de función en este gen en más de un 20% de LLA-T<sup>426,429</sup>. No obstante, mediante secuenciación con una elevada profundidad de lectura y PCR tipo *droplet*, alcanzando sensibilidades  $10^{-3}$ - $10^{-4}$ , no se ha logrado identificar mutaciones en este gen en el diagnóstico<sup>428</sup>.

Otro gen de recaída, implicado en el metabolismo de fármacos, es ***NR3C1*** pues codifica para un receptor de glucocorticoides. En los estudios realizados, principalmente en cohortes pediátricas de LLA-T, no es un gen frecuentemente mutado al diagnóstico, pero si en la recaída<sup>222,430</sup>. Se han evaluado funcionalmente mutaciones concretas que afectan a este gen, mostrando que confieren resistencia a prednisolona, glucocorticoide utilizado en el tratamiento de la LLA-T<sup>430</sup>.

El gen supresor de tumores ***TP53*** es el segundo gen más comúnmente mutado en las recaídas de la LLA-T, encontrándose alterado en aproximadamente un 25%<sup>288,316,317</sup>. En este caso es posible encontrar alteraciones en ***TP53*** en muestras al diagnóstico, pero en una frecuencia menor que en la recaída. Las mutaciones de ***TP53*** están asociadas a un alto riesgo de fracaso al tratamiento y mal pronóstico<sup>288,316,317</sup>.

Las mutaciones presentes en la vía de señalización RAS-MAPK no son exclusivas de la recaída y su patrón de evolución clonal es altamente heterogéneo, pues en algunos casos se observan mutaciones comunes al diagnóstico y la recaída, pero en otras ocasiones se pierden/ganan en la recaída respecto al diagnóstico<sup>417,431</sup>. Este hecho, sugiere que estas mutaciones son frecuentemente subclonales y que la presión de la quimioterapia puede resultar en una selección positiva o negativa<sup>222</sup>. Estas incluyen mutaciones activadoras en ***KRAS***, ***NRAS*** y ***PTPN11***<sup>222,426</sup>. Estudios funcionales han mostrado que estas mutaciones generan resistencia a glucocorticoides y metotrexato<sup>222,290,432</sup>.

Las alteraciones en reguladores epigenéticos también se han visto asociadas con la recaída de la LLA. Es el caso de las alteraciones que afectan a ***CREBBP***, ***KMT2D***, ***MLL2***, ***SET2D***, ***SUZ12***, ***WHSC1*** y ***SMARCA4***<sup>14</sup>. La presencia de mutaciones en la mayoría de estos genes

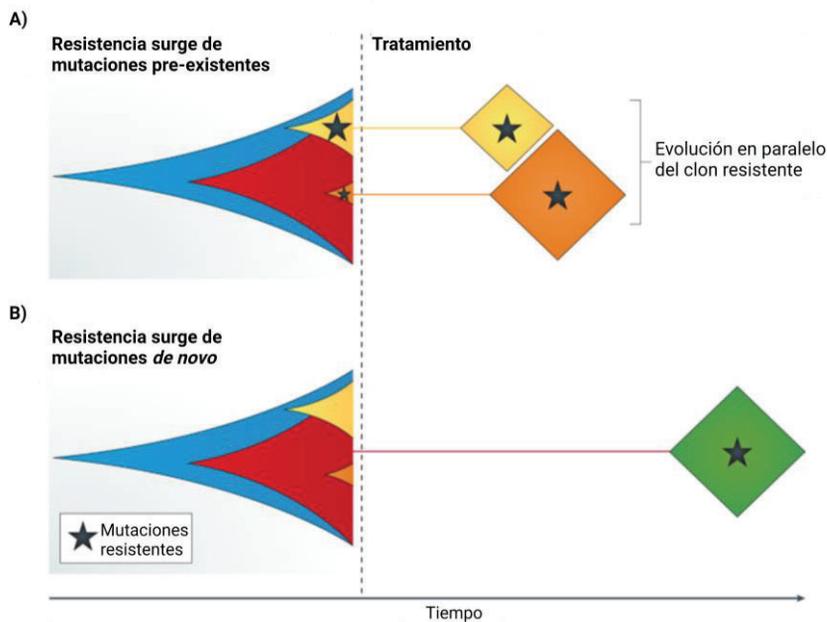
también se ha asociado a resistencia a glucocorticoides, en concreto, las mutaciones en *CREBBP*<sup>433</sup>, *WHSC1*<sup>434</sup> y *SET2D*<sup>435</sup>.

El perfil mutacional característico de la recaída muestra que, a diferencia de las mutaciones comunes del diagnóstico, que afectan a genes implicados en la leucemogénesis, los genes mutados en la recaída están involucrados en generar resistencia a los fármacos utilizados en el tratamiento de la LLA-T o facilitar el escape a una respuesta de lesión del ADN, entre otros mecanismos. Además, cabe destacar que el perfil mutacional al diagnóstico de la LLA-T y LLA-B es diferente. Sin embargo, ambas neoplasias siguen un esquema de tratamiento similar y tienen un perfil mutacional en la recaída similar<sup>436</sup>.

Por tanto, junto al modelo de evolución clonal anteriormente explicado (Fig. 9), el perfil mutacional característico de la recaída sugiere que los clones responsables de la recaída ya estarían presentes en el momento del diagnóstico y en la mayoría de los casos, en un clon ancestral. La quimioterapia sería la principal fuerza selectiva ejercida, seleccionando aquellos clones con mutaciones que les conferirían una ventaja selectiva (Fig. 10A).

No obstante, estos resultados se contraponen a otra de las principales teorías del origen de la recaída, la cual indica que hay mutaciones responsables de su origen, pero en este caso estas se habrían originado por el efecto de la propia quimioterapia (Fig. 10B). Hay diferentes evidencias que apoyan esta teoría: i) las mutaciones en el gen *NTC52* se han encontrado exclusivamente en la recaída y no se identifican en el diagnóstico incluso a elevadas sensibilidades<sup>414,428</sup>; ii) en otros tipos de cánceres se ha observado que la quimioterapia genera unas determinadas marcas sobre las células debido a las lesiones que se producen en el ADN<sup>437-439</sup>; iii) en el caso de la LLA hay un estudio en pacientes pediátricos donde se observó que el tratamiento con tiopurinas generaba una nueva firma mutacional y que algunas de las mutaciones presentes, de forma exclusiva en la recaída (*PRPS1*, *NT5C2* y *TP53*), se podrían haber generado por el efecto de la quimioterapia<sup>430</sup>. En este mismo estudio, se utilizó un modelo que estimó el tiempo de duplicación de los blastos y se observó que no era compatible con el hecho de que las mutaciones responsables de la recaída estuvieran ya presentes en clones anteriores al diagnóstico, sugiriendo que el tratamiento era el principal responsable de la adquisición de las nuevas variantes<sup>430</sup>. No obstante, este modelo fue cuestionado posteriormente, poniendo en evidencia su validez<sup>440</sup>. Además, en otro trabajo posterior donde se utilizaron modelos computacionales para estimar el punto exacto de divergencia entre los clones del diagnóstico y la recaída, se mostró que en la mayoría de los casos no había pasado suficiente tiempo entre ambos puntos para que la leucemia encontrada en la recaída hubiese evolucionado a partir de una alteración adquirida durante el tratamiento y que por tanto, tendría que haber evolucionado a partir de un clon anterior a la quimioterapia<sup>415</sup>.

Así pues, existe un gran debate sobre el origen de los clones responsables de la recaída, siendo una limitación la insuficiente sensibilidad de las técnicas actualmente disponibles.



**Figura 10. Principales teorías sobre el origen de la recaída.** Adaptada de Turajlic et al., 2019<sup>441</sup>

### 3.7.2.2 Factores pronósticos de la recaída

La persistencia o reaparición de ERM positiva después de la quimioterapia de inducción/consolidación es el factor pronóstico de recaída más importante actualmente<sup>378,442-444</sup>. No obstante, existe un alto porcentaje de pacientes que a pesar de tener una ERM negativa terminan sufriendo una recaída. Además, ésta no informa de resistencias específicas a tratamientos sin dar opción a probar tratamientos dirigidos a dianas concretas.

La identificación de alteraciones características de la recaída en el momento del diagnóstico sería un factor clave para prevenir una recaída y poder administrar esquemas de tratamiento diferentes que no generasen resistencias. El trabajo publicado por Trinquand *et al.*<sup>263</sup> ha sido el único en mostrar la presencia de alteraciones genéticas asociadas a una incidencia de recaída en la LLA-T. En concreto, los pacientes con mutaciones en *NOTCH1/FBXW7* y sin mutaciones en *N/KRAS/P TEN* presentaban menor incidencia de recaída que el resto de pacientes<sup>263</sup>. No obstante, estos resultados no fueron replicados posteriormente<sup>375</sup>. Por tanto, en la actualidad no existen factores genéticos pronósticos de recaída validados.

### 3.7.2.3 Nuevos agentes para pacientes recaídos/refractarios

Actualmente, las opciones de tratamiento de recaída en la LLA-T son muy limitadas. El TPH es la aproximación más utilizada junto a la repetición de pautas de quimioterapia distintas a las usadas en el diagnóstico, aunque las tasas de remisiones conseguidas son bajas<sup>445</sup>. En este sentido, se plantea el uso de tratamientos con mecanismos de acción diferentes, especialmente aquellos que van dirigidos contra dianas que se sabe que son resistentes a

los bloques de quimioterapia actuales. Como se muestra en la figura 11 existen fármacos para la mayoría de las alteraciones características de la LLA-T. A continuación, se resumen los principales:

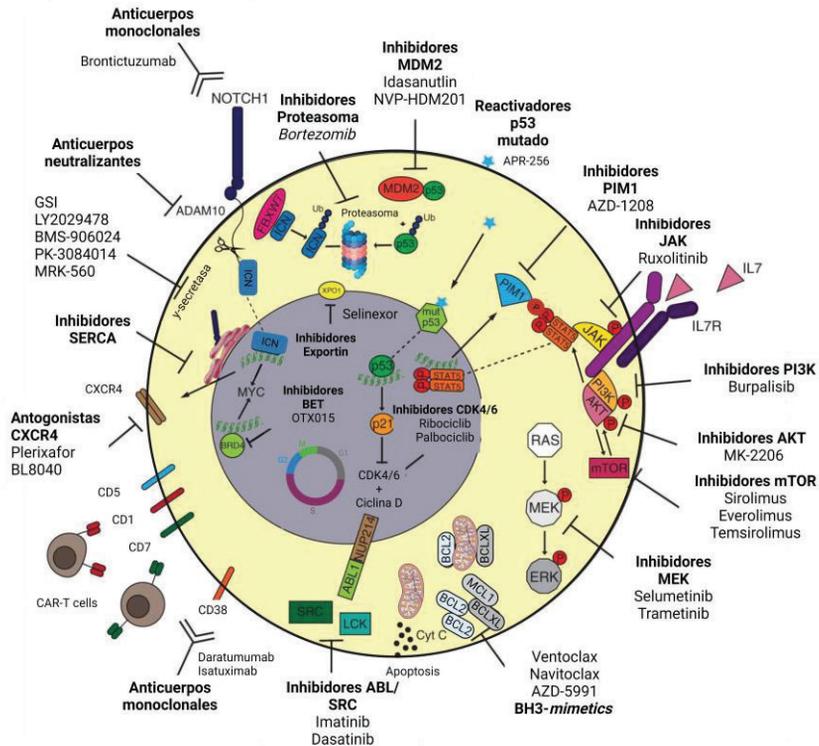


Figura 11. Terapias dirigidas en al LLA-T. Adaptada de Cordo' *et al.*, 2021 <sup>446</sup>

- **Análogos de purinas: Nelarabina**

Es un análogo del nucleósido purina que inhibe la síntesis del ADN y que ha mostrado eficacia en el tratamiento de la LLA-T. Se ha visto que la acumulación por parte de los linfoblastos T del compuesto ara-G, derivado del metabolismo de la nelarabina, tiene un efecto citotóxico<sup>447</sup>. En la actualidad, la nelarabina es el único fármaco aprobado para el tratamiento de pacientes recaídos de LLA-T/LBL, teniendo unas tasas de respuesta del 50% para pacientes pediátricos y adultos jóvenes<sup>448</sup>, mientras que en adultos son un poco más bajas (31% y 36% alcanzan la remisión completa en dos estudios independientes), pero aun así muestra beneficios para este grupo de pacientes ya que puede inducir remisiones clínicas que facilitan un ulterior TPH <sup>449,450</sup>.

- **Modificadores de la apoptosis: BH3 Mimetics**

Los compuestos de tipos BH3 *mimetics* son péptidos que se unen a las proteínas anti-apoptóticas, facilitando la apoptosis de las células del cancer<sup>451</sup>. Este tipo de aproximación

ha supuesto una revolución pues se puede cuantificar la sensibilidad que las células van a tener hacia los agentes BH3 *mimetics* conociendo la expresión de las proteínas apoptóticas en el blasto. Además, la respuesta mitocondrial a la apoptosis se correlaciona con la respuesta clínica de la LLA y predice la quimio-sensibilidad, lo que permite hacer un cribado de agentes BH3 *mimetics*<sup>452</sup>. Venetoclax es un fármaco de tipo BH3 *mimetics* antagonista de la proteína anti-apoptótica BCL2 aprobado por la FDA en el tratamiento de la leucemia linfática crónica con resultados muy favorables<sup>453</sup>. Las LLA-T más inmaduras expresan altos niveles de BCL-2 que van disminuyendo con la maduración al contrario de los que ocurre con BCL-XL<sup>454,455</sup>. Es por ello, que se ha propuesto el uso de venetoclax para el tratamiento de las LLA-T más inmaduras mientras que navitoclax, compuesto antagonista de BCL-XL, para las LLA-T más maduras. Los primeros estudios en pacientes pediátricos y adultos recaídos o refractarios tratados con venetoclax muestran resultados muy prometedores<sup>456,457</sup>. Actualmente hay en curso ensayos clínicos que combinan venetoclax y navitoclax para el tratamiento de la LLA, tanto B como T, refractaria o en recaída.

- **Inhibidores principales vías de señalización activadas en la LLA-T**

Inhibidores de NOTCH1: La activación constitutiva del receptor mutado de NOTCH1 requiere de la acción de una gamma-secretasa. Es por ello, que se han propuesto inhibidores de gamma-secretasa en el tratamiento de la LLA-T. No obstante, no han llegado a la clínica debido a su toxicidad gastrointestinal y tener una limitada actividad anti-leucémica<sup>458</sup>.

Una estrategia alternativa es la inhibición más específica de la actividad de las gamma-secretasas mediante el uso de inhibidores de su región catalítica PSEN1 (MRK-560) con resultados pre-clínicos muy prometedores<sup>459</sup>.

Otras estrategias propuestas para bloquear la señalización alterada de NOTCH1 incluye el tratamiento con anticuerpos monoclonales<sup>460</sup>, inhibidores que bloquean la maduración de la proteína NOTCH1 evitando su posicionamiento en la membrana celular<sup>461</sup> o inhibir la acción de proteínas promovidas por la activación de la vía de NOTCH1, como CXCR4<sup>462</sup>.

Inhibidores de IL7R-JAK-STAT: Son varios los inhibidores que se han diseñado contra esta vía, mostrando eficacia en modelos pre-clínicos<sup>463,464</sup>. Ruxolitinib es un inhibidor de JAK1/2 aprobado por la Administración de Alimentos y Medicamentos (en inglés Food and Drug Administration, FDA) y utilizado en el tratamiento de neoplasias mieloproliferativas, que bloquea la señalización de JAK-STAT a pesar de la presencia de mutaciones<sup>465</sup>. Se ha visto que ruxolitinib muestra eficacia y que actúa de manera sinérgica junto a la dexametasona en LLA-T y ETP-ALL con resistencia a esteroides inducida por IL-7<sup>466</sup>.

Inhibidores de PI3K-AKT-mTOR: Se han diseñado inhibidores PIK3 tipo pan<sup>292,467-469</sup>, inhibidores duales PI3K/mTOR<sup>470,471</sup>, inhibidores duales mTORC1/mTORC2<sup>472</sup>, y el inhibidor directo de AKT (MK-2206)<sup>467,473</sup>, que bloqueen la activación ectópica de esta vía de señalización.

Inhibidores de MEK: Inducen la muerte de células con mutaciones en *RAS* y tienen un efecto sinérgico con glucocorticoides en células primarias de pacientes con LLA-T. Destaca

selumetinib que está en ensayos clínicos en combinación con dexametasona para el tratamiento de pacientes adultos y pediátricos recaídos con LLA-B y LLA-T <sup>291</sup>.

- **Inhibidores ciclo celular**

Inhibidores de CDK: *CDKN2A/B* son reguladores negativos de CDK4/6 y es posible inhibir la acción de estas ciclinas mediante dos agentes palbociclib y ribociclib, mostrando resultados positivos en estudios preclínicos <sup>474,475</sup>.

Por otro lado, la acción de las deleciones en *CDKN1B* que llevan a la activación de CDK2 puede ser bloqueada por inhibidores de CDK1/2/5/9, como dibaciclib <sup>476</sup> compuesto que ha mostrado eficacia en estudios en fase preclínica aunque no se ha conseguido demostrar en ensayos clínicos <sup>477</sup>.

Fármacos dirigidos contra TP53: Estos fármacos restauran la actividad normal de p53 y se encuentran en fases preclínicas de desarrollo <sup>478</sup>. En concreto, el reactivador APR-246 ha mostrado efectos positivos en cultivos ex vivo de un paciente recaído tras ser trasplantado <sup>479</sup> y en pacientes con mutaciones en *TP53* afectados por otras neoplasias hematológicas <sup>478</sup>.

- **Inhibidores de la Familia quinasas Src (ABL1)**

Las células con el reordenamiento *NUP214::ABL1* responden al tratamiento con TKI utilizados en el tratamiento de la leucemia mielocítica crónica (LMC) y LLA Ph<sup>+326</sup> y algunos casos, a pesar de no presentar alteraciones en *ABL1*, son sensibles a los tratamientos con imatinib y dasatinib en modelos preclínicos <sup>480</sup>. Además, también se ha visto que las LLA-T con elevada expresión de LCK se pueden beneficiar de dasatinib <sup>481,482</sup>. Por ello, se resalta la importancia de estudiar el perfil de fosforilación de las proteínas como biomarcadores de respuesta a inhibidores de la familia SRC <sup>483</sup>.

- **Reguladores epigenéticos**

Inhibidores de BET: La familia de proteínas BET regulan la transcripción génica mediante la acetilación de histonas <sup>339</sup>. Hay un tipo de estas proteínas (BRD4) que promueve la expresión aberrante de *MYC* en la leucemia y dado que las LLA-T con *NOTCH1* mutado tienen un incremento en la expresión de *MYC* se ha propuesto un inhibidor de BRD4 (JQ1) como tratamiento dirigido a pacientes recaídos <sup>484</sup>. El tratamiento con JQ1 ha mostrado una reducción del tumor en modelos preclínicos <sup>484</sup>. Aunque JQ1 ha mostrado una corta vida media, estos hallazgos han abierto una puerta al desarrollo de diferentes inhibidores de BET <sup>485</sup>. Las LLA-T MYC positivas han mostrado sensibilidad a los inhibidores de BET en modelos preclínicos <sup>172,486</sup>.

Inhibidores de HDAC: Las histonas deacetilasas (HDAC) son enzimas clave en la remodelación de la cromatina y en la regulación epigenética de los genes. Los inhibidores de HDAC panobinostat, vorinostat y romidepsin se están utilizando en el tratamiento de mieloma y linfomas. Específicamente, panobinostat ha mostrado eficacia en modelos in vivo de LLA-T mejorando la supervivencia en combinación con quimioterapia <sup>487</sup>.

Inhibidores de metiltransferasas: La decitabina y la azacitidina son dos inhibidores de metiltransferasas que inducen la hipometilación de la cromatina provocando una alteración en la transcripción génica. Ambos están aprobados para el tratamiento de SMD y están en fases tempranas de ensayos clínicos para pacientes pediátricos de LLA. Un estudio ha demostrado que el pretratamiento con decitabina potencia el efecto de la quimioterapia en modelos preclínicos de ETP-ALL<sup>488</sup> y ha mostrado resultados prometedores en el tratamiento de un paciente adulto recaído con ETP-ALL<sup>489</sup>.

- **Terapia basada en anticuerpos monoclonales**

Los anticuerpos anti-CD38 son los principalmente propuestos para el tratamiento de la LLA. CD38 es un receptor transmembrana que se expresa en linajes mieloides, linfoides y también en algunas células no hematológicas. Se expresa en la LLA-T y en las ETP-ALL durante toda la progresión de la leucemia<sup>490</sup>. Daratumumab es un anticuerpo monoclonal anti-CD38 aprobado por la FDA y por la Agencia Europea de Medicamentos (en inglés *European Medicines Agency*, EMA) para el tratamiento del mieloma múltiple en pacientes recaídos/refractarios. CD38 y ha presentado una alta eficacia en modelos PDX de LLA-T<sup>490</sup>.

- **Terapia celular: CAR-T**

Los CAR (receptor de antígeno quimérico, en inglés *Chimeric Antigen Receptor*) son polipéptidos de origen sintético formados por un dominio extracelular (cadena ligera de inmunoglobulina con el fragmento variable) dirigidos contra un antígeno determinado; una región *hinge* o bisagra; un dominio transmembrana y un dominio de señalización intracelular. Los CAR-T son linfocitos T modificados genéticamente para que expresen constitutivamente el receptor CAR contra el antígeno determinado. Esta expresión se consigue mediante la infección de los linfocitos T por lentivirus o retrovirus que integran la secuencia genética que codifica para el CAR deseado. El resultado es la generación de linfocitos citotóxicos contra células que expresen el antígeno determinado(Fig. 12)<sup>491</sup>.

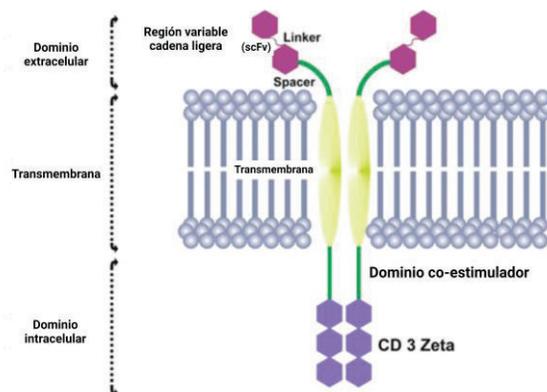


Figura 12. Estructura CAR. Adaptada de Sheykhhasan *et al.*,2022<sup>491</sup>.

El diseño de CAR-Ts para tratar la LLA-T supone un gran reto pues se tienen que conseguir suficientes células T del paciente sin presencia de blastos. Además, se puede dar lo que se conoce como efecto fratricida, y es que las células CAR-T comparten los mismos marcadores de superficie que los blastos contra los que van dirigidas, lo que autoextinguiría las células CAR-T. A pesar de estas limitaciones, se han diseñado varios CAR-T para enfermedades de células T, dirigidos contra CD5, CD7, CD1a y CD38.

Una de las dianas más atractivas es CD7, ya que se expresa en aproximadamente el 95% de las LLA-T. Se han utilizado diferentes estrategias para evitar el efecto fratricida, como por ejemplo, la eliminación del gen CD7 de las células CAR-T mediante CRISPR-Cas9<sup>492</sup>. No obstante, dado que CD7 se expresa en todos los timocitos y células T, los pacientes que reciben este tratamiento pueden sufrir una eliminación de este tipo celular y una inmunodeficiencia. Una solución a esto es diseñar CAR-T que contienen “genes suicidas” inducibles y que pueden ser activados selectivamente<sup>493</sup>, o utilizar un CAR-T fratricida-resistente desarrollado a partir de células T de un donante sano<sup>494</sup>. En estos últimos, se modifican genéticamente para hacerlas deficientes en CD7 y eliminar el gen *TCRAD* para eliminar el riesgo de enfermedad injerto contra huésped. También se están desarrollando CAR-T anti CD7 resistentes al fratricidio al bloquear el receptor T, sin necesidad de eliminación genética.

Recientemente se ha desarrollado un CAR-T fratricida-resistente anti-CD1a para tratar LLA-T CD1a positivas<sup>495</sup>. CD1a se expresa solo durante la fase proliferativa del desarrollo de los timocitos y no durante la fases inmaduras y maduras, lo que limita el riesgo de inmunodeficiencia completa

- **Otras terapias**

XPO1 es una proteína exportadora nuclear-citoplasmática involucrada en el transporte de varias proteínas que participan en la regulación del ciclo celular (p53 y pRB) y está *up-regulado* en varias neoplasias incluidas la LLA-T. Selinexor es un inhibidor selectivo de la exportación nuclear y un antagonista de XPO1 aprobado por la FDA para el tratamiento del mieloma. Se ha demostrado en fases preclínicas la eficacia de selinexor en el tratamiento de la LLA-T, aunque ha generado toxicidades<sup>496</sup>.

Un método diferente de alterar las vías de señalización oncogénicas es mediante la regulación de la estabilidad o degradación de las proteínas implicadas en estas. Bortezomib es un inhibidor del proteasoma, lo que impide la degradación de las proteínas provocando un bloqueo del ciclo celular y facilitando la apoptosis. Este compuesto está aprobado en el tratamiento del mieloma múltiple refractario y un estudio ha demostrado resultados positivos en el tratamiento de la LLA-T<sup>497</sup>. Otra alternativa, es mediante la inhibición de la enzima NAE que se encarga de ubiquitinar y regula la actividad e interacciones entre proteínas esenciales implicadas en la regulación del ciclo celular<sup>498</sup>. Datos preclínicos han mostrado que un inhibidor de esta enzima, pevonedistat, puede inducir el bloqueo del ciclo celular y la apoptosis en modelos de LLA-T<sup>498</sup>. Tanto bortezomib como pevonedistat están actualmente en investigación clínica para pacientes con LLA.





# HIPÓTESIS Y OBJETIVOS



La hipótesis global de esta tesis es que la caracterización genética de pacientes adultos de LLA-T tratados de un modo homogéneo ayuda a predecir la respuesta al tratamiento en estos pacientes y la progresión de la leucemia hacia la recaída.

Teniendo en cuenta lo anteriormente expuesto, los objetivos de esta Tesis Doctoral son:

---

### 1.

Profundizar en la comprensión de la implicación de las alteraciones genéticas en el desarrollo de la LLA-T mediante la caracterización genómica de una cohorte retrospectiva de pacientes en el momento del diagnóstico:

- 1.1 Identificar las alteraciones genéticas, SNV/indels y CNV, involucradas en el desarrollo de la neoplasia
  - 1.2 Evaluar la cooperación de las diferentes alteraciones genéticas en el desarrollo de la LLA-T y su involucración en el proceso leucemogénico y en las características biológicas de los pacientes.
  - 1.3 Estudiar la composición y jerarquía clonal de los pacientes adultos con LLA-T.
- 

### 2.

Evaluar el impacto clínico de las alteraciones genéticas encontradas al diagnóstico en una cohorte de pacientes adultos con LLA-T tratados de manera homogénea y re-estratificar a los pacientes en función de los marcadores pronósticos identificados.

---

### 3.

Ampliar el conocimiento de la repercusión de las alteraciones genéticas en el proceso de recaída de pacientes adultos con LLA-T:

- 3.1 Identificar alteraciones genéticas responsables de la recaída y con valor predictivo
- 3.2 Ampliar la comprensión sobre el origen de la recaída.





# METODOLOGÍA



## METODOLOGÍA DEL OBJETIVO 1

### 1. Pacientes y diseño del estudio

Para el primer objetivo, se analizaron mediante TDS un total de 146 muestras retrospectivas de pacientes con de LLA-T diagnosticados en 39 hospitales de España. En 135 de ellas también se analizó las alteraciones tipo CNV. El diagnóstico de la LLA-T se realizó de acuerdo con los criterios de la OMS<sup>88</sup> y la clasificación citogenética en base al estudio de Genescà *et al*<sup>374</sup>.

Se obtuvo el consentimiento informado de todos los pacientes incluidos en el estudio. Las muestras y los datos clínicos se almacenaron de acuerdo con la declaración de Helsinki. Este estudio fue aprobado por la *Institutional Review Board* del Hospital Germans Trias i Pujol. En la tabla 4 se resumen las características biológicas de los pacientes de esta cohorte.

**Tabla 4. Características clínico-biológicas pacientes cohorte Objetivo 1**

		n=146
<b>Edad, años</b>	Mediana [extremos]	36 [3 - 86]
<b>Género</b>	Hombre	108 (74%)
	Mujer	38 (26%)
<b>Citogenética</b>	Nº alt<3	85/141 (60%)
	Nºalt≥3	12/141 (8,5%)
	NE (no evaluable)	44/141 (31,2%)
	ND (no disponible)	5 (3,4%)
<b>Inmunofenotipo</b>	ETP-ALL	31/138 (22,5%)
	Pre-T	28/138 (20,3%)
	Tímica cortical	54/138 (39,1%)
	Tímica madura	25/138 (18,1%)
	No clasificable	8 (5,5%)
<b>Protocolo de tratamiento</b>	Pediátrico	3 (2%)
	LAL-RI-08	3 (2%)
	LAL-OLD/FRAIL-07	6 (4%)
	AR-03	32 (22%)
	AR-11	86 (59%)
	LAL19	15 (10%)
	ND	1 (0,7%)

Nº alt: Número de alteraciones; NE: No evaluable; ND: No disponible

#### 1.1. Procesamiento de las muestras, extracción del ADN y control de calidad

El material de partida para realizar los estudios genómicos fue ADN genómico procedente de células blásticas de MO, o en su defecto de SP, almacenado a -20°C.

### 1.1.1. Descongelación

En los casos en los que no se disponía de ADN, este se obtuvo de células blásticas congeladas, provenientes de la propia colección del grupo de investigación ALL almacenadas en el Biobanco del IGTP (Instituto Germans Trias i Pujol, Badalona) (PT17/0015/0045), del Banco Español de ADN Carlos III (PT13/0001/0037), del Biobanco de La Fe (PT13/0010/0026) y de colecciones de diferentes hospitales pertenecientes al grupo PETHEMA. Los viales se encontraban congelados en nitrógeno líquido a  $-196^{\circ}\text{C}$ . Se descongelaron parcialmente al baño María a  $37^{\circ}\text{C}$  y a continuación se añadió una solución que incluyó medio *Hibernate™-A* (cat. no. A1247501 Thermo Fisher, Massachusetts, EE. UU.) + 10% FBS (suero bovino fetal, del inglés *Fetal Bovine Serum*) (cat. no. 10270, Thermo Fisher, Massachusetts, EE. UU.) pre-calentado a  $37^{\circ}\text{C}$ . Por cada mL de muestra incluida en el vial se añadió de manera gradual un volumen determinado de medio: primero 1 mL seguido de una incubación de un minuto, a continuación 2 mL de medio seguido de una incubación de un minuto, 5 mL de medio seguido de una incubación de 1 minuto y finalmente 5 mL de medio. Las células junto al medio se centrifugaron durante 5 minutos a  $300 \times g$ . Tras la centrifugación se retiró el sobrenadante, resuspendiendo las células en PBS (tampón fosfato salino, del inglés *Phosphate buffered saline*) 1X. En caso de que fuese necesario, las células se contaron usando la cámara de Neubauer.

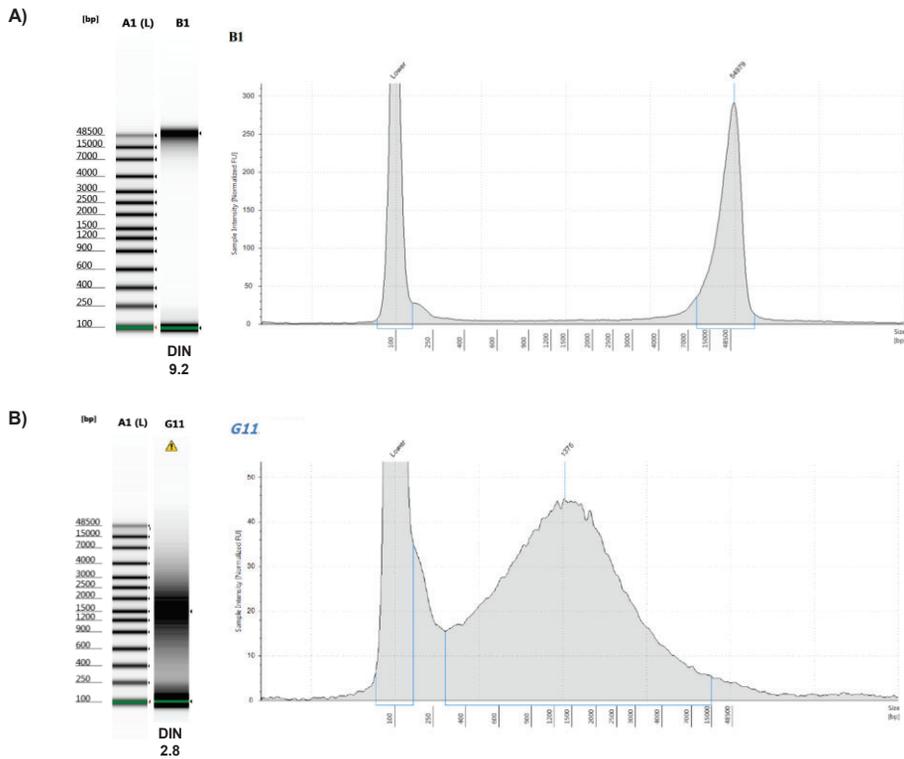
### 1.1.2. Enriquecimiento de las muestras en contenido de blastos T (*sorting*)

Las muestras con un porcentaje de blastos menor al 80% fueron sometidas a un proceso de *sorting*, seleccionando una población CD45+ intermedia y CD7+, para enriquecerlas en blastos tipo T (*FACSAria II*, BD Biosciences, New Jersey, EE. UU.). Para ello, las células descongeladas se marcaron con los anticuerpos anti-CD45 (APC Mouse Anti-Human CD45) (cat. no. 555485) y anti-CD7 (PE Mouse Anti-Human CD7) (cat. no. 555361) (BD Biosciences, Nueva Jersey, EE. UU.), añadiendo 1:5 ul de anticuerpo por cada millón de células. La incubación con los anticuerpos se realizó durante 30 minutos a temperatura ambiente, y posteriormente, se hizo un lavado con PBS 1X, resuspendiendo las células en volumen final de 2,5 mL de PBS 1X + 10% FBS.

### 1.1.3. Extracción del ADN y control de calidad

El ADN se extrajo usando el *kit Maxwell RSC Cultured Cells* (cat. no. AS1620, Promega, Wisconsin, EE. UU.) siguiendo las instrucciones del fabricante. La cuantificación del ADN se realizó mediante el espectrofotómetro *NanoDrop™ One/OneC* (Thermo Fisher Scientific, Massachusetts, EE. UU.) que además de informar de la concentración del ADN, permite conocer la presencia de contaminantes en la muestra. Se evaluaron los *ratios* 260/280 nm, que dan información de la presencia de proteínas en la muestra, y 260/230, que permiten evaluar la presencia de otros contaminantes como sales, solventes orgánicos u otros contaminantes ambientales. Para las técnicas que lo requirieron, el ADN también se cuantificó usando el reactivo *Qubit 3.0 high sensitivity* (HS, cat. no. Q32851) */broad range* (BR, cat. no. Q32850, Invitrogen Life Technologies, California, EE. UU.), que cuantifica ADN de doble cadena. Se evaluó la integridad del ADN (en inglés *DNA integrity number*, DIN) mediante la plataforma de electroforesis *TapeStation* (cat. no. G2991BA, Agilent Technologies, California, EE. UU.), concretamente usando los reactivos para ADN genómico,

(cat. no. 5067-5366, Agilent Technologies, California, EE. UU.). En la figura 13 se muestra un ejemplo del perfil obtenido del análisis con la *TapeStation*.



**Figura 13. Integridad ADN evaluada con TapeStation.** A) Perfil de ADN íntegro. B) Perfil de ADN fragmentado.

## 1.2. Diseño del panel de secuenciación y validación

### 1.2.1. Diseño del panel de secuenciación

Se diseñó un panel de secuenciación con Agilent Technologies (Santa Clara, CA). Para el diseño de las sondas se realizó una revisión bibliográfica con el fin de incluir todas aquellas alteraciones que estaban descritas por tener una implicación en la LAL-T. La primera recopilación bibliográfica y diseño de regiones a incluir en el panel, fue realizada por Jesús García-Chico, estudiante de Máster del grupo en 2018.

Las regiones que se incluyeron fueron las siguientes:

- Exones completos de 85 genes con mutaciones descritas al diagnóstico de la LAL-T. Supusieron un total 352 exones y una longitud de 176408 pb.
- Regiones promotoras de 15 genes codificantes de reguladores transcripcionales de la LLA-T. Se incluyeron regiones situadas 2000 pares de bases aguas arriba (*upstream*) del lugar de inicio de la transcripción con la finalidad de incluir las regiones de las cajas TATA y CAAT. Estas regiones supusieron un total de 16409 pb.

- 5 potenciales regiones que podrían actuar como *enhancers* de la LAL-T. Mediante el análisis de la base de datos PrESSTo, que incluye regiones descritas como *enhancers*, se seleccionaron aquellas que estaban expresadas en timocitos y que afectaban a genes de interés de la LLA-T. Estas regiones supusieron un total de 3597 pb.
- Además, también se incluyeron regiones con mutaciones relacionadas con la recaída, con el fin de que este panel de genes pudiese utilizarse como una herramienta predictiva de la recaída. Por ello, se incluyeron exones completos de 4 genes con mutaciones descritas en la recaída de la LLA-T. Supusieron un total de 21 exones y una longitud de 5047 pb. Además, se tenía la información del análisis mutacional de 15 muestras pareadas al diagnóstico y recaída de pacientes con LLA-T y se incluyeron exones completos de 15 genes con mutaciones solo identificadas en las muestras de la recaída. Estos últimos supusieron un total de 5047 pb.

El tamaño final del panel de genes diseñado fue de 206042 pb. En la figura 14 se muestran los genes para los cuales se incluyó alguna región en el panel.

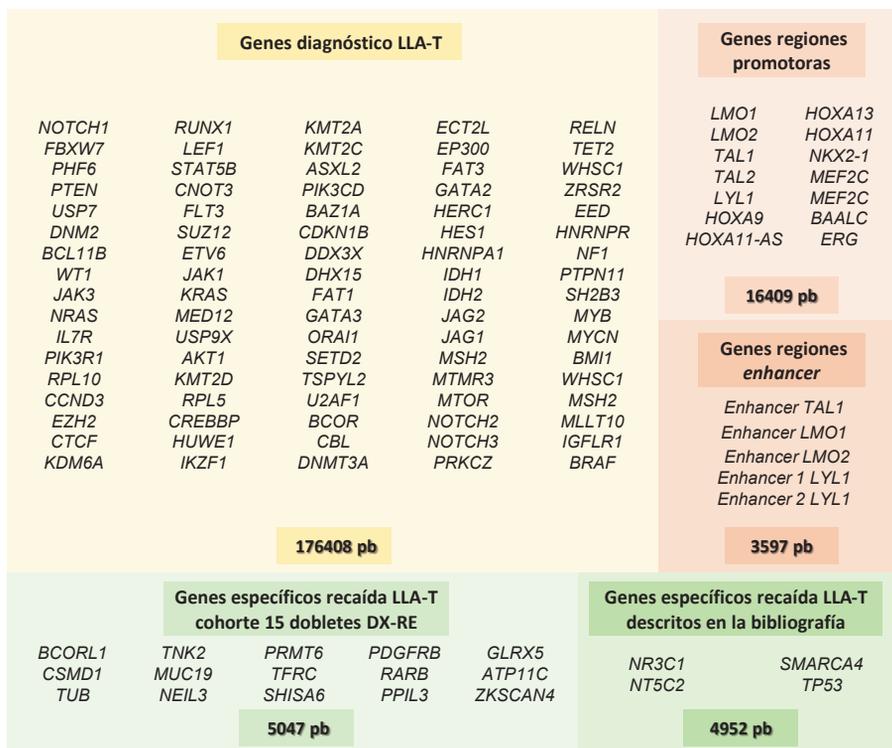


Figura 14. Regiones incluidas en el panel de genes de LLA-T

Este panel fue diseñado con dos objetivos. En primer lugar, para utilizarlo con fines de investigación, para caracterizar genéticamente los pacientes incluidos en este estudio y poder evaluar el impacto pronóstico de las alteraciones halladas. Además, para utilizarlo como herramienta diagnóstica. Actualmente se utiliza en la rutina diagnóstica para analizar el perfil mutacional de los genes recomendados por el protocolo de tratamiento LAL2019 [NCT04179929] de PETHEMA.

### 1.2.2. Validación

Con el fin de validar el diseño del panel y la técnica, se secuenciaron, con el panel, 10 muestras de las cuales disponíamos de los datos de WGS. El proceso de validación consistió en cuatro pasos:

1. Evaluación de las regiones cubiertas: haciendo uso de un visor genómico (*Integrative Genomics Viewer*, IGV) se comprobó que todas las regiones incluidas en el diseño habían sido secuenciadas en las muestras de validación.
2. Estudio de la profundidad de lectura: para cada una de las regiones incluidas en el diseño, se evaluó el número de veces que se había leído y aquellas regiones que no tenían la profundidad adecuada, se modificó el diseño añadiendo más sondas para llegar a las profundidades deseadas.
3. Especificidad: se cuantificó la cantidad de lecturas que no se correspondieron con regiones incluidas en el diseño (*off target*) y se optimizó el protocolo de preparación de las librerías para reducirlas, haciendo especial énfasis en los pasos de los lavados tras la hibridación con las sondas, con el fin de eliminar todas aquellas secuencias no hibridadas.
4. Reproducibilidad: Se compararon las variantes detectadas mediante TDS con aquellas detectadas anteriormente mediante WGS con el fin de comprobar que todas las variantes situadas en regiones incluidas en el diseño detectadas por WGS también se secuenciaron utilizando el panel diseñado. En la Tabla 5 se muestran las variantes identificadas con cada plataforma y se observa que todas fueron secuenciadas por ambas plataformas, con la excepción de una variante de *NOTCH1* solo detectada por TDS. Esta discordancia podría ser debido a la diferencia de profundidad de lectura de cada una de las aproximaciones (90X mediante WGS frente a 240X mediante TDS). Además, se realizó un control externo cruzado: se secuenció una muestra usando otro panel de genes específico de LLA diseñado con Illumina (California, EE. UU.), obteniendo una homología del 100%

**Tabla 5. Comparación variantes detectadas por WGS frente a las detectadas por TDS**

ID	GEN	TRANSCRITO	MUTACIÓN	VAF (%)	
				WGS	Panel NGS
PAT1	<i>IKZF1</i>	NM_006060	p.K136fs	47,73	29,31
	<i>IKZF1</i>	NM_006060	p.K136T	52,33	35,29
	<i>NOTCH1</i>	NM_017617	p.K1607delinsFPK	17,5	19,05
	<i>NOTCH1</i>	NM_017617	p.F1606L	16,25	21,77
	<i>KRAS</i>	NM_033360	p.A18D	49,46	41,67
	<i>KMT2D</i>	NM_003482	p.P647Q	32,56	52,17
	<i>PHF6</i>	NM_032458	p.C20fs	69,44	60

	<i>NOTCH1</i>	NM_017617	p.L1574P	7,79	3,97
	<i>NOTCH1</i>	NM_017617	p.L1585P	-	7,04
PAT2	<i>NRAS</i>	NM_002524	p.G12D	57,89	49,39
	<i>DNMT3A</i>	NM_022552	p.G543A	98,15	98,39
PAT5	<i>DNMT3A</i>	NM_022552	p.D876E	44,12	41,18
	<i>DNMT3A</i>	NM_022552	p.I715fs	43,36	45,39
	<i>TNK2</i>	NM_005781	p.P162fs	45,22	46,26
	<i>TFRC</i>	NM_003234	p.T229I	55,77	46,06
	<i>IL7R</i>	NM_002185	p.243_245del	46,34	19,38
	<i>IL7R</i>	NM_002185	p.I245M	52,7	16,32
	<i>IL7R</i>	NM_002185	p.S246G	53,95	18,56
	<i>NOTCH1</i>	NM_017617	p.T2466fs	44,05	37,3
	<i>NOTCH1</i>	NM_017617	p.F1592delinsFYS	39,86	41,94
	<i>SUZ12</i>	NM_015355	p.R286X	45,26	42,86
	<i>JAK3</i>	NM_000215	p.R657Q	47,66	42,29
	<i>EP300</i>	NM_001429	p.Q2288R	48,57	51,78
	<i>PHF6</i>	NM_032458	p.R274X	45,38	45,83
PAT7	<i>NOTCH1</i>	NM_017617	p.L1574P	48,78	39,85
	<i>PHF6</i>	NM_032458	p.T255fs	87,88	80,39
PAT8	<i>NOTCH1</i>	NM_017617	p.A1701P	12,5	21,75
	<i>CDKN1B</i>	NM_004064	p.E172X	35,38	38,78
	<i>BCL11B</i>	NM_138576	p.L55fs	32,86	20,23
	<i>NOTCH1</i>	NM_017617	p.L1678P	8,57	7,4
PAT9	<i>NOTCH1</i>	NM_017617	p.V2443fs	38,57	28,75
	<i>NOTCH1</i>	NM_017617	p.V2443L	42,86	32,86
	<i>IGFLR1</i>	NM_024660	p.L250fs	50,79	46,15
	<i>RUNX1</i>	NM_001754	p.R166delinsRDX	27,27	22,22
	<i>PHF6</i>	NM_032458	p.S154X	34,21	26,32
PAT10	<i>NOTCH1</i>	NM_017617	p.L1593P	29,92	34,56
	<i>FAT3</i>	NM_001008781	p.R2460Q	55,86	50,54
	<i>PHF6</i>	NM_032458	p.306_310del	32,98	32,39
	<i>NOTCH1</i>	NM_017617	p.V1599A	7,75	9,15
	<i>NOTCH1</i>	NM_017617	p.L1596H	8,4	8,84
	<i>JAK3</i>	NM_000215	p.E958delinsDG	9,92	9,53
	<i>DNM2</i>	NM_001005361	p.Q239X	4,8	4,63
PAT11	<i>LEF1</i>	NM_016269	p.N140fs	52,38	45,58
	<i>FBXW7</i>	NM_033632	p.R465H	49,46	47,65
	<i>NSD2</i>	NM_133330	p.Q103R	48,57	46,05
PAT12	<i>NRAS</i>	NM_002524	p.G12V	41,74	41,84
	<i>PTEN</i>	NM_000314	p.T232fs	34,4	30,4
	<i>ETV6</i>	NM_001987	p.I209fs	93,94	80,85
	<i>PHF6</i>	NM_032458	p.R319X	45,13	23,08
PAT15	<i>NOTCH1</i>	NM_017617	p.D2108N	41,81	40,7
	<i>NOTCH1</i>	NM_017617	p.F1592S	49,12	45,68

VAF: *variant allele frequency* (frecuencia alélica de la variante); WGS: *whole genoma sequencing* (secuenciación del genoma completo); NGS: *next generation sequencing* (secuenciación de nueva generación).

### 1.3. Detección y análisis de las alteraciones tipo SNV/indel

#### 1.3.1. Preparación de las librerías, enriquecimiento por hibridación con sondas y captura

El material de partida fue ADN genómico. Los criterios de calidad del ADN fueron:

- Presencia de contaminantes evaluada por *NanoDrop*:
  - *Ratio* de absorbancia a 260/280 nm: 1,8-2
  - *Ratio* de absorbancia a 260/230 nm: ~2
- Cuantificación por *Qubit 3.0 BR*: 75-150 ng de ADN
- Integridad evaluada por *TapeStation*: *DIN* > 7

El ADN fue fragmentado mediante un *kit* comercial de enzimas de restricción (*SureSelect XT HS Enzymatic Fragmentation Kit*, cat. no. 5191-4079, Agilent Technologies, California, EE. UU.). Se evaluó el perfil de fragmentación mediante la plataforma de electroforesis *TapeStation* usando los reactivos para DNA de baja concentración (*4200 TapeStation D1000 HS*, cat. no. 5067-5585, Agilent Technologies, California, EE. UU.).

Posteriormente, se prepararon y amplificaron las librerías usando el *kit SureSelect XT HS* (cat. no. 5500-0138, Agilent Technologies, California, EE. UU.) y se evaluó el perfil de las librerías mediante *TapeStation*, usando los reactivos *4200 TapeStation D1000* (cat. no. 5067-5583, Agilent Technologies, California, EE. UU.).

A continuación, se realizó la hibridación con las sondas de ARN biotiniladas, previamente diseñadas para enriquecer el ADN en las regiones de interés. Tras la hibridación, el ADN se precipitó utilizando bolas unidas a estreptavidina (cat. no. 65602, Invitrogen, California, EE. UU.) que se unen con afinidad a la biotina de las sondas. El ADN se lavó, se eluyó y se amplificó usando índices únicos para cada muestra.

Tras este proceso las librerías definitivas se cuantificaron cualitativa y cuantitativamente usando el *Qubit 3.0 HS* (cat. no. Q32851, Invitrogen Life Technologies, California, EE. UU.) para conocer la concentración y la *TapeStation* para evaluar el tamaño de los insertos de la librería. En la figura 15 se muestra un esquema de este proceso.

#### 1.3.2. Secuenciación

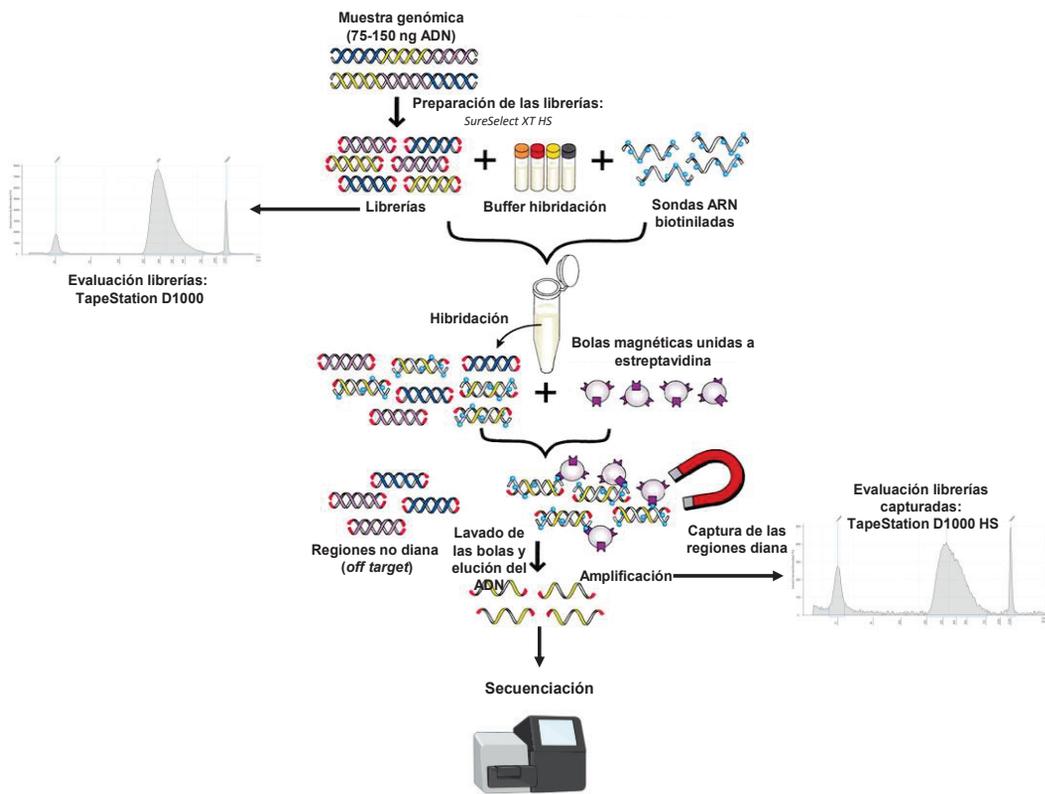
Las librerías fueron secuenciadas en una plataforma *MiSeq* (Illumina) del servicio de genómica del IGTP, siguiendo un método *paired-end*, con un tamaño de lecturas de 2 x 75 pb. Tras el proceso de secuenciación la plataforma recoge la información de las secuencias de todas las muestras en un fichero tipo FASTQ.

#### 1.3.3. Procesamiento de los datos secuenciados

Los ficheros FASTQ pasaron un primer control de calidad con el programa *FastQc* (v0.11.8). Posteriormente se evaluó el *Phred quality score*, un parámetro logarítmico que representa la fiabilidad con que se identifica cada base secuenciada. En el caso de este trabajo, solo se tuvieron en cuenta las secuenciadas con un *Phred quality score* superior a 20, es decir,

secuencias con una probabilidad de error menor al 1%. Estas secuencias fueron alineadas al genoma de referencia (versión hg19) haciendo uso de la herramienta BWA (v0.7.151).

Las lecturas mapeadas fueron recalibradas usando GATK (v3.4.462). Los duplicados de PCR fueron enmascarados mediante la herramienta Picard (v1.1383). El llamamiento de las variantes (*variant calling*) se realizó mediante la combinación de SamTools (v1.104) y VarScan2 (v2.4.05). Posteriormente, se anotaron los resultados usando ANNOVAR (v2018-04-166). El procesamiento de los datos secuenciados fue realizado por el técnico bioinformático del grupo.



**Figura 15. Diagrama del protocolo de preparación de las librerías y enriquecimiento del ADN** en las regiones de interés por hibridación con sondas y captura. Adaptada de página oficial Agilent Technologies.

### 1.3.4. Filtrado de las variantes

Para el filtrado de variantes se realizaron un conjunto de pasos secuenciales:

1. Filtrado de variantes polimórficas: Se eliminaron del análisis las variantes descritas en  $>0,1\%$  de las bases de datos poblacionales 1000Genomes, ExAC, gnomAD, y Exome Variant Server.
2. Filtrado de las variantes intrónicas y sinónimas. Se seleccionaron únicamente las variantes exónicas no sinónimas, pudiendo ser de diferentes tipos: variantes con cambio

de sentido (*missense*), variantes sin sentido (*nonsense*), pequeñas deleciones/inserciones (indels) que afectaron a la pauta de lectura (*frameshift*), o que no afectaron a la pauta de lectura (*non-frameshift*).

3. Filtrado manual de las variantes utilizando el visor genómico IGV (*Integrative Viewer Genomics*) para eliminar posibles artefactos de la secuenciación. Se eliminaron las variantes: leídas solo en una dirección (*strand bias*), incluidas en regiones repetitivas, con una profundidad de lectura no uniforme y localizadas en el extremo de las lecturas (*edge effect*).
4. Filtrado de las variantes que no cumplieron los criterios de profundidad de lectura establecidos: variantes con una profundidad total de la posición de <30 X; variantes leídas en <8 lecturas, variantes con una frecuencia alélica (en inglés *variant allele frequency*, VAF) < 5%. No obstante, hubo variantes que cumplieron los criterios de profundidad, pero tuvieron una VAF < 5% que se mantuvieron en el análisis por ser variantes descritas por su implicación en la LLA-T (*hotspot*). Fue el caso de variantes: *NOTCH1*\_NM\_017617 (p.L1680S, p.L1574P, p.L1678P, p.Q2405X, p.T1573delinsKS, p.V1578E, p.V1676D, p.V1721E) *JAK1*\_NM\_002227 (p.R724C, p.R724H, p.Y652H), *PTEN*\_NM\_000314 (p.E288fs, p.T232fs); *JAK3*\_NM\_000215 (p.E958K); *FBXW7*\_NM\_033632 (p.R465H); *BCL11B*\_NM\_138576 (p.R447C); *EP300*\_NM\_001429 (p.N2053H).
5. Clasificación de las variantes según su impacto funcional potencial: se evaluaron un total de 10 predictores *in-silico* provistos por el paquete de anotación ANNOVAR: SIFT, Polyphen2, LRT, MutationTaster, MutationAssessor, FATHMM, PROVEAN, VEST3, MetaSVM, MetaLRD. Con el fin de validar su valor predictivo, se partió de variantes reportadas y validadas funcionalmente con impacto directo en la proteína y en la LLA-T y se observó que era necesario que más de 5 predictores asignasen a una variante como patogénica para considerarla como tal. A partir de ahí, las variantes se categorizaron en 3 grupos según su potencial impacto funcional:
  - Benignas: Variantes no reportadas en las bases de datos ClinVar<sup>499</sup> y COSMIC<sup>500</sup> y >5 predictores *in-silico* la marcan como benigna
  - Variantes de significado incierto (en inglés *Variants unknown significance*, VUS): las bases de datos ClinVar<sup>499</sup> y COSMIC<sup>500</sup> no dan información del impacto de la variante ni tampoco los predictores, o no es concluyente.
  - Patogénicas: Variantes reportadas en las bases de datos ClinVar<sup>499</sup> y COSMIC<sup>500</sup> y/o >5 predictores la marcan como patogénica.

En esta tesis únicamente se contemplaron las variantes anotadas como VUS y patogénicas.

Cabe decir que sin un control germinal de cada muestra no se garantiza la eliminación de todas las variables de origen germinal. Por tanto, después de este filtrado podría quedar alguna variante germinal.

En la figura 16 se muestra un resumen del análisis y filtrado de las secuencias.

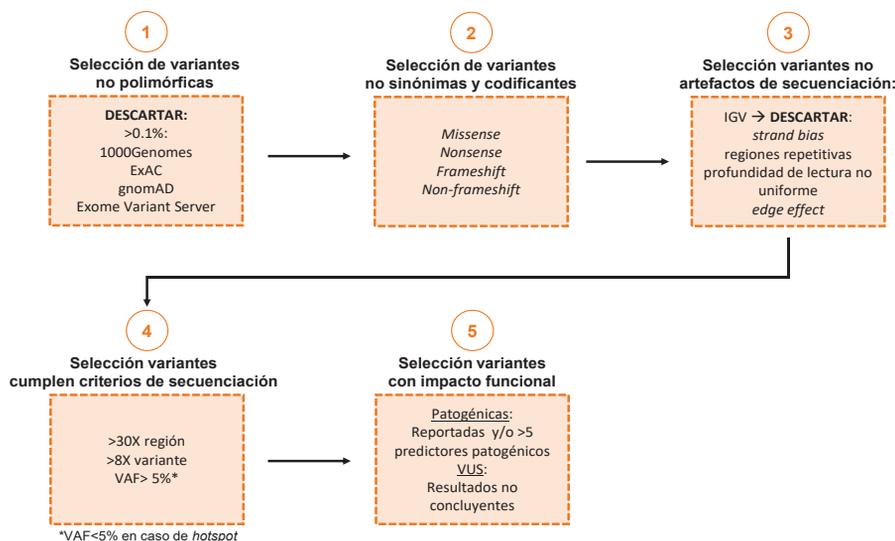


Figura 16. Diagrama del filtrado de las alteraciones detectadas por NGS

### 1.3.5. Comparación del perfil mutacional de muestras pareadas sangre periférica y médula ósea

Se secuenciaron 15 muestras pareadas de MO y SP (Tabla 6) y se comparó el perfil mutacional para confirmar que ambos tipos de muestras aportaban la misma información genética y, por tanto, se podían utilizar de manera indistinta en los estudios genómicos. Aquellas muestras que se disponía de células y tuvieron <80% blastos se sortearon para reducir diferencias en el contenido de blastos. Todas las frecuencias alélicas de las variantes fueron corregidas siguiendo el método explicado en el apartado 1.5.

Tabla 6. Cohorte seleccionada estudio MO-SP

ID	MO		SP	
	%blastos	Profundidad de lectura media (X)	%blastos	Profundidad de lectura media (X)
PAT1	94	307	84	315
PAT2	95	292	90	450
PAT3	95	411	90	238
PAT4	90	315	85	217
PAT5	97	389	100*	223
PAT6	100	331	67	386
PAT7	76	281	100*	298
PAT8	95	269	100*	269
PAT9	100	331	100*	249
PAT10	89	512	83	264
PAT11	95	78	100*	496
PAT12	94	385	100*	277
PAT13	95	477	80	335
PAT14	93	490	100*	1815
PAT15	68	567	100*	303

\*Muestras sorteadas



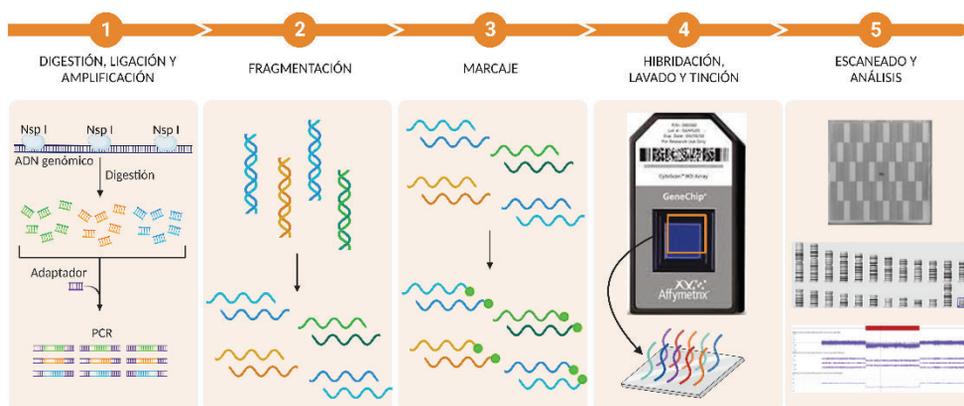
de 25 nucleótidos y el espacio medio entre cada sonda es de 2 kb para el caso de las sondas CNV y de 200 Mb para el caso de las sondas SNP. Cada SNP es detectado por 6 sondas, 3 para cada alelo.

Esta técnica fue realizada por la Unidad de Microarrays del Instituto Josep Carreras.

El material de partida fueron 550 ng de ADN genómico. Los criterios de calidad del ADN fueron:

- Presencia de contaminantes evaluada por *NanoDrop*:
  - *Ratio* de absorbancia a 260/280 nm: 1,8-2
  - *Ratio* de absorbancia a 260/230 nm: ~2
- Cuantificación por *NanoDrop*: 50 ng/  $\mu$ l
- Integridad evaluada por *TapeStation*: *DIN* > 7

Brevemente, para realizar esta técnica el ADN se digiere con una enzima específica (*Nsp*I), se amplifica, se marca con nucleótidos terminadores marcadores con biotina y se hibrida al *array*. Posteriormente, el *array* se lava y tiñe y se escanea para detectar los cambios de fluorescencia. En la figura 18 se representa un esquema de este proceso.



**Figura 18. Esquema del proceso de detección de alteraciones CNV por SNP-array.**

El análisis de los resultados se realizó utilizando el *software Chromosome Analysis Suite* (ChAS v 4.0.0.385) provisto por el fabricante (Thermo Fisher Scientific, Massachusetts, EE. UU.). El filtrado de las variantes detectadas consistió en un conjunto de pasos secuenciales:

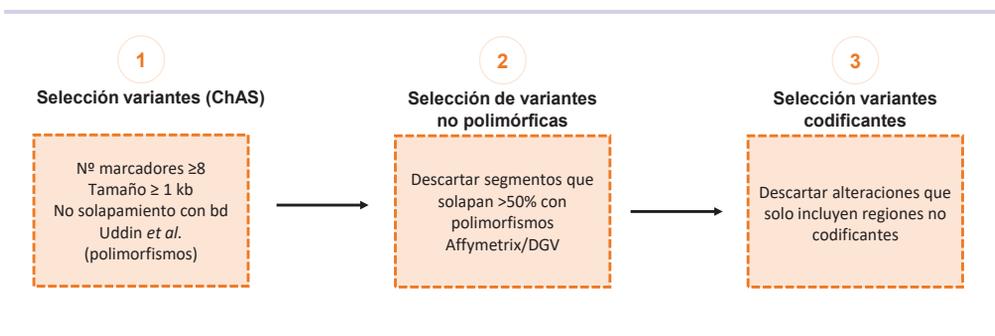
1. Filtrado de las variantes utilizando los ajustes del programa ChAS:
  - Número de marcadores que cubren la alteración: Se estableció que para considerar una alteración como tal tenía que estar cubierta al menos con 8 marcadores.
  - Tamaño de la alteración: Se filtraron aquellas alteraciones menores a 1 kpb.
  - Solapamiento con variantes detectadas en bases de datos poblacionales: Se estableció como base de datos de variantes polimórficas el estudio realizado por Uddin *et al.*<sup>501</sup>

donde analizaron un total de 873 controles sanos utilizando la misma plataforma (*CytoScan HD*, Thermo Fisher Scientific) y los mismos criterios de tamaño (1kb) y número de sondas positivas (>8 sondas) para un evento. Mediante estudios de validación se determinó que aquellas alteraciones que solapaban >90% con alteraciones detectadas al menos en un paciente del estudio de Uddin *et al.*<sup>501</sup> eran polimórficas. Con este parámetro se filtraron la mayoría de los polimorfismos.

2. Filtrado de polimorfismos comparando las CNV con los registrados en la base de datos poblacional DGV (*Database of Genomic Variants*) y con los datos derivados del análisis de 2421 controles sanos provistos por Thermo Fisher Scientific. En este caso, se filtraron las variantes que solapaban al menos un 50% con alguna de las variantes reportadas en estas bases.
3. Filtrado de todas las variantes en regiones no codificantes.

Las deleciones correspondientes a los reordenamientos somáticos del gen del receptor de las células T (TCR) y de los genes de las inmunoglobulinas fueron excluidos del análisis.

En la figura 19 se muestra un resumen del filtrado y análisis de las CNV.



**Figura 19. Diagrama del filtrado de las alteraciones detectadas por SNP-array.**

### 1.5. Corrección de la frecuencia alélica y estimación de la jerarquía clonal.

Partiendo de la base de que las mutaciones se adquieren de manera secuencial, según el modelo de la teoría clonal de la evolución del cáncer, aquellas mutaciones adquiridas más tarde en el tiempo afectarán a un menor número de células que las mutaciones que aparecen en el clon fundador, que estarán en todas las células<sup>502</sup>. Por tanto, para estimar esta evolución se estimó la proporción de células afectadas por cada alteración.

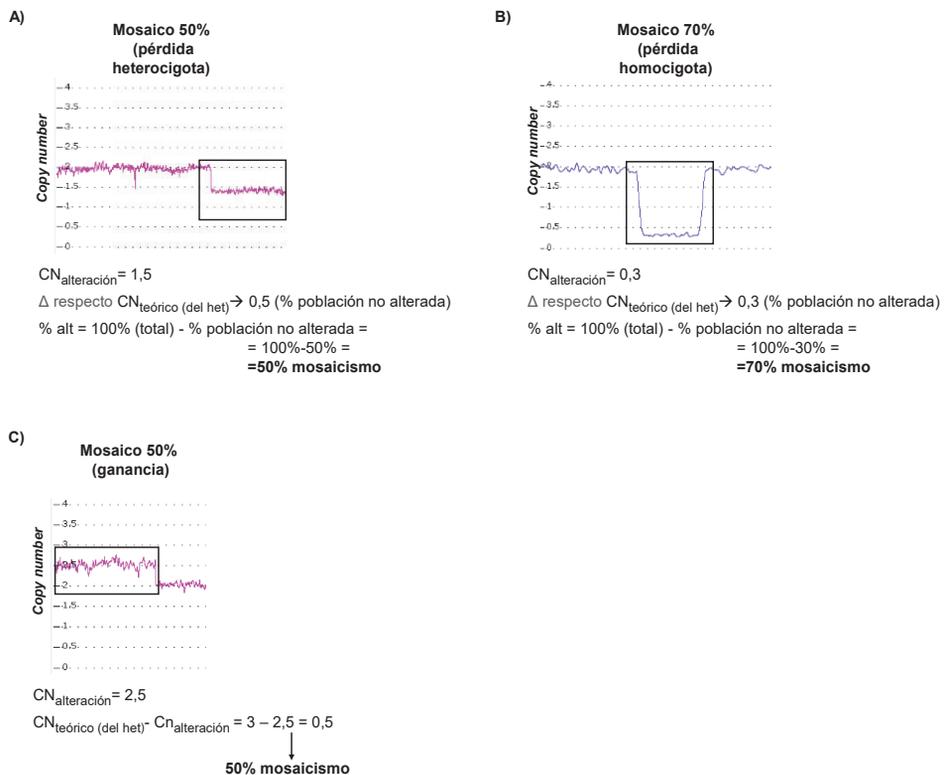
En el caso de las alteraciones tipo SNV/indel, partiendo de la VAF de cada una de ellas, se puede inferir las mutaciones que pertenecen a cada clon.

Para el caso de las alteraciones tipo CNV el programa de análisis no proporciona el valor de la VAF, pero este se corresponde con el mosaicismo de las muestras, y este último se puede inferir a partir del análisis de dos parámetros: *copy number* y *allele difference*.

*Copy number*: este valor toma un número entero en el caso de que la muestra analizada no tenga mezcla de poblaciones (mosaicismo). De este modo, CN=0 indica deleción

homocigota, CN=1 indica delección heterocigota y CN=3 indica ganancia de una copia. En aquellos casos en los que este valor no sea un número entero es debido a la presencia de mezcla poblaciones. En función de estos valores se estimó el % de la muestra con la alteración (“VAF”):

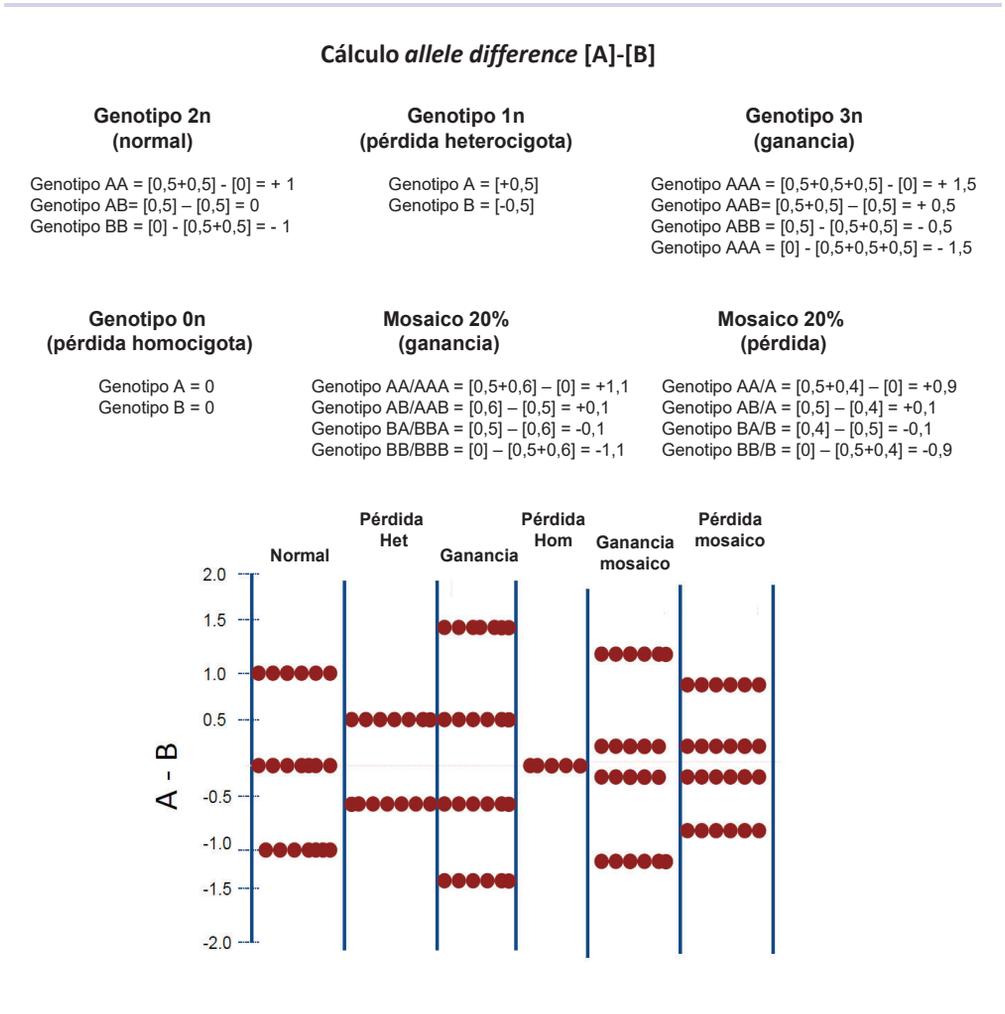
- Pérdidas heterocigotas: valor del CN para alteraciones en mosaicismo es  $1 < CN < 2$ . Lo que aumenta respecto a 1 indica la proporción de células que no tienen la alteración. En la figura 20A se detalla cómo se infirió el % de la alteración.
- Pérdidas homocigotas: valor del CN para alteraciones en mosaicismo es  $0 < CN < 1$ . Lo que aumenta respecto a 0 indica la proporción de células que no tienen la alteración. En la figura 20B se detalla cómo se infirió el % de la alteración.
- Ganancias: el valor del CN para alteraciones en mosaicismo es  $2 < CN < 3$  (en el caso de la ganancia de una copia). La diferencia respecto al CN teórico de la ganancia (CN=3) indica la proporción de células que no tiene la alteración. En la figura 20C se detalla cómo se infirió el % de la alteración. Esta aproximación es similar cuando la ganancia es de  $>3$  copias (CN=4, CN=5, CN=6...).



**Figura 20. Detección de mosaicismo en SNP-arrays (Sondas copy number)** A) Ejemplo pérdida heterocigota en mosaico 50%. B) Ejemplo pérdida homocigota en mosaicismo 70%. C) Ejemplo ganancia (1 copia) en mosaicismo 50%.

El *allele difference* es un parámetro obtenido a partir del análisis de las sondas SNPs. Para obtener este parámetro el programa de análisis asigna cada alelo del SNP como alelo A y alelo B, tomando el alelo A un valor de 0,5 y el alelo B un valor de -0,5. El *allele difference* se estima como la diferencia entre  $[A] - [B]$ .

De este modo, en función del tipo de alteración el *allele difference* toma diferentes valores mostrados en la figura 21, donde también se muestra cómo se visualizan estos valores en caso de población homogénea y en caso de presencia de mosaicismo.



**Figura 21. Valores y representación del parámetro *allele difference* para los diferentes tipos de alteraciones CNV.**

Las VAF de todas las alteraciones (SNV/indel y CNV) se convirtieron en la fracción celular de la variante, es decir, la fracción celular del cáncer (en inglés *cancer cell fraction*, CCF).

Para ello, se tuvo en cuenta que la VAF depende de dos factores: la pureza de la muestra, es decir, la proporción de células leucémicas del total de células de la muestra; y del número de copias que tenga la región estudiada, esta última solo para el caso de las alteraciones tipo SNV/indel.

La ploidía del genoma también influye, pero en todos los casos analizados de LLA-T fue normal (2n). Tampoco hubo ningún caso donde hubiese una mutación y duplicación en la misma posición.

Corrección por la pureza y *copy number*.

$$CCF = \frac{VAF \times copy\ number}{pureza} \times 100$$

La CCF adquiere valores de 0 a 1 considerándose cercano a 1 cuando afecta a la práctica totalidad de las células leucémicas.

En esta tesis se consideraron subclonales aquellas alteraciones que tuvieron una CCF<50% (VAF<25%), mientras que se consideraron clonales aquellas alteraciones que tuvieron una CCF>50% (VAF>25%).

## 1.6. Asociaciones entre mutaciones y características biológicas

Se evaluó la presencia de asociaciones entre parejas de genes alterados en los mismos pacientes como también asociaciones de genes alterados con determinadas características biológicas del paciente.

Esto se realizó en 135 pacientes de LLA-T (número de pacientes de los que se disponía de datos de secuenciación y análisis de CNV) y se tuvieron en cuenta los genes que se encontraron mutados en  $\geq 10$  pacientes.

En concreto, las características biológicas estudiadas fueron: edad, inmunofenotipo y cariotipo. Los pacientes fueron clasificados en cuatro grupos de edad: <18 años (n=7), 18-35 años (n=59), 36-60 años (n=58) y >60 años (n=8); por subtipo de LLA-T según el inmunofenotipo: ETP-ALL (n=29), pre-T (25), cortical (n=48) y madura (n=25); y por número de alteraciones presentes en el cariotipo: cariotipo con menos de 3 alteraciones ( $n^{\circ} \text{alt} < 3$ ) (n=77), cariotipo con al menos 3 alteraciones ( $n^{\circ} \text{alt} \geq 3$ ) (n=12) y cariotipo no evaluable (NE) (n=44).

Las asociaciones entre las diferentes parejas se calcularon usando la prueba exacta de Fisher, posteriormente corregido con la prueba *Benjamini-Hochberg*, asumiendo una significancia de  $q < 0,05$ .

## METODOLOGÍA DEL OBJETIVO 2

### 1. Pacientes y diseño del estudio

Para el segundo objetivo, la cohorte de estudio retrospectiva estuvo compuesta por 116 pacientes para el caso del análisis de las alteraciones SNV/indels y por 109 para el caso del análisis de las alteraciones tipo CNV. Todos ellos estuvieron incluidos en la cohorte retrospectiva analizada en el Objetivo 1. El resto de los pacientes fueron descartados por no cumplir los criterios de inclusión requeridos para este estudio:

- Pacientes adultos con una edad comprendida entre 16 y 61 años.
- Pacientes tratados según dos protocolos de LLA alto riesgo consecutivos y guiados según la respuesta de la EMR: LAL-AR/2003[Ph][NCT00853008]<sup>381</sup>, LAL-AR/2011[Ph][NCT01540812]<sup>380</sup>.
- Pacientes de los que se disponía de datos de la ERM al final de la inducción-1 (alrededor del día 35 desde el inicio del tratamiento).

En la tabla 7 se resumen las características clínico-biológicas de la cohorte de este estudio.

**Tabla 7. Características clínico-biológicas pacientes cohorte Objetivo 2**

		n=116
Edad, años	Mediana [extremos]	37 [16 - 61]
Género	Hombre	86 (74%)
	Mujer	30 (26%)
ECOG	0	43/112 (38%)
	1	53/112 (47%)
	2	14/112 (13%)
	≥3	2/112 (2%)
Recuento WBC (x10 <sup>9</sup> /L)	Mediana [extremos]	50,5 [0,5 - 525,4]
Adenopatías	No	43/98 (44%)
	Sí	55/98 (56%)
Esplenomegalia	No	72/111 (65%)
	Sí	39/111 (35%)
Hepatomegalia	No	85/110 (77%)
	Sí	25/110 (23%)
Masa mediastínica	No	64/113 (57%)
	Sí	49/113 (43%)
Infiltración del SNC	No	96/110 (87%)
	Sí	14/110 (13%)
Citogenética	Nº alt<3	66 (57%)
	Nºalt≥3	10 (9%)
	NE	40 (34%)
Inmunofenotipo	ETP-ALL	22 (19%)
	Pre-T	19 (16%)
	Tímica madura	22 (19%)
	Tímica cortical	47 (41%)
	No clasificable	6 (5%)

ECOG: *Eastern Cooperative Oncology Group* (Grupo cooperativo de oncología del Este); WBC: *whole blood cells* (recuento de glóbulos blancos); SNC: sistema nervioso central; Nº alt: Número de alteraciones; NE: no evaluable

## 2. Esquemas de tratamiento

Todos los pacientes incluidos en este estudio fueron tratados con protocolos diseñados por el grupo PETHEMA. Concretamente los protocolos AR-03 (Fig. 22) y AR-11 (Fig. 23). Se trata de protocolos guiados según la respuesta de la ERM. Brevemente:

- AR-03: Los pacientes se trataron con quimioterapia de inducción (vincristina, prednisona y daunorubicina, principalmente; acompañados del tratamiento triple intratecal [TIT]). El porcentaje de blastos se evaluó a día +14 por examen morfológico y de acuerdo con este valor los pacientes continuaron con un bloque de inducción estándar (< 10% blastos) o una inducción de intensificación (>10%). Tras el tratamiento de inducción, los pacientes en RC siguieron 3 bloques de consolidación estándar (dexametasona, metotrexato, Ara-C, L-asparaginasa y mercaptopurina; acompañados de TIT) y los pacientes que no alcanzaron la RC siguieron con un único bloque de consolidación. La respuesta al bloque de consolidación se evaluó en función de la ERM. Aquellos pacientes con EMR < 0,05% continuaron con una quimioterapia de mantenimiento (mercaptopurina y metotrexato) y reinducciones estándar (vincristina y prednisona). Los pacientes con EMR > 0,05% fueron candidatos a un TPH junto a aquellos que tuvieron > 10% blastos al día +14.
- AR-11: Los pacientes realizaron una pre-fase al tratamiento estándar de inducción (prednisona acompañada de TIT), seguida de la fase de inducción (vincristina, daunorubicina, prednisona, asparaginasa, TIT). Al final de esta fase (~+35 días) se evaluaron los niveles de ERM y los pacientes se estratificaron de acuerdo con estos valores: EMR < 0,1% siguieron un tratamiento de consolidación estándar (3 bloques de consolidación: vincristina, dexametasona, metotrexato/ Ara-C, asparaginasa; acompañados de TIT), EMR > 0,1% siguieron un tratamiento de intensificación de inducción (FLAG-IDA) seguido por un TPH. Los pacientes con EMR < 0,01% tras el bloque de consolidación continuaron con una quimioterapia de mantenimiento (mercaptopurina y metotrexato) y reinducciones (vincristina, prednisona, TIT).

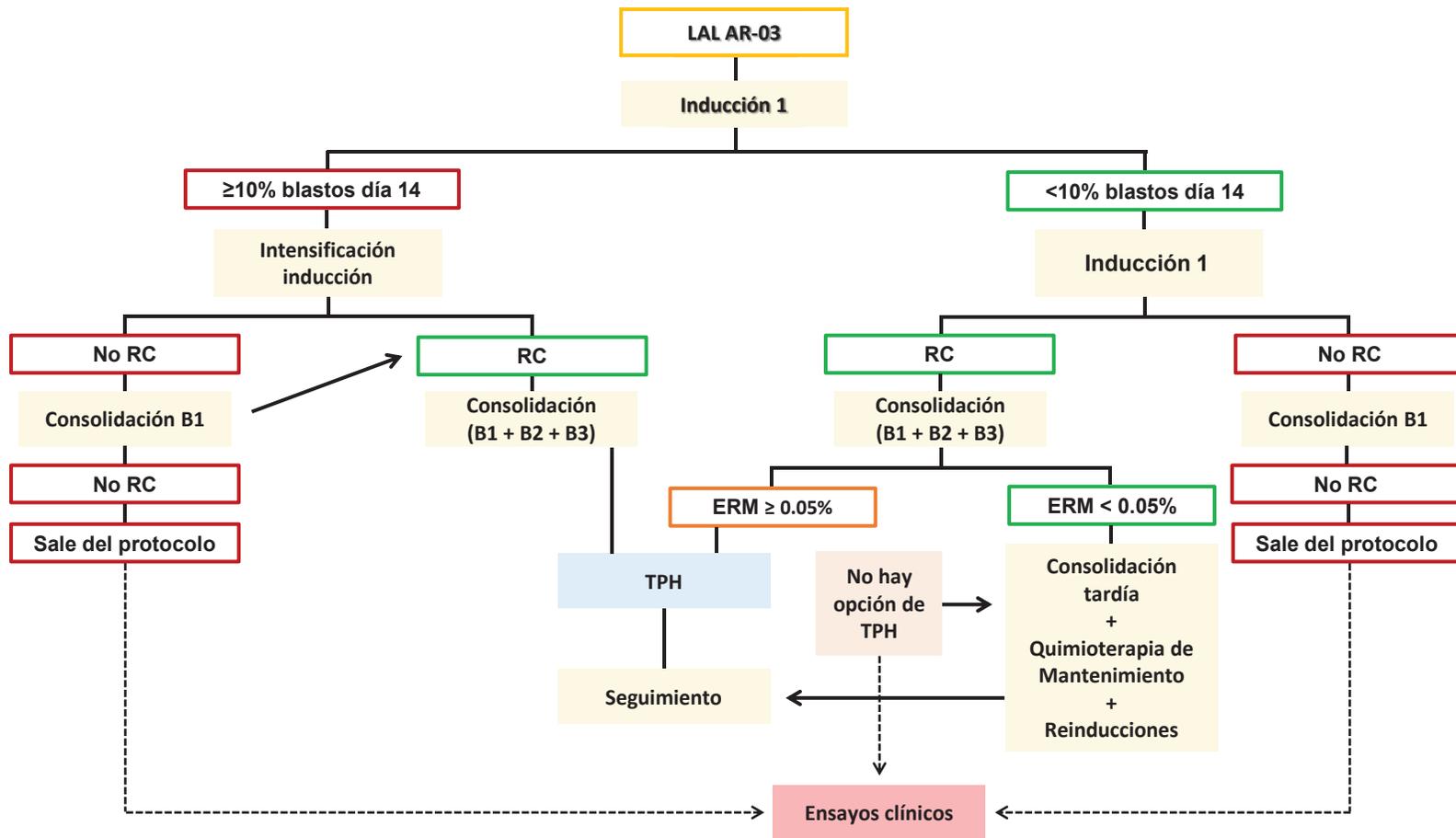
## 3. Análisis estadísticos

Las características clínicas y biológicas de la serie se describieron como frecuencia y porcentajes para las variables categóricas y como mediana y extremos para las variables cuantitativas. La prueba de Chi-cuadrado o el *test* exacto de Fisher se utilizó para comparar proporciones, y la prueba de la mediana para comparar las medianas entre grupos.

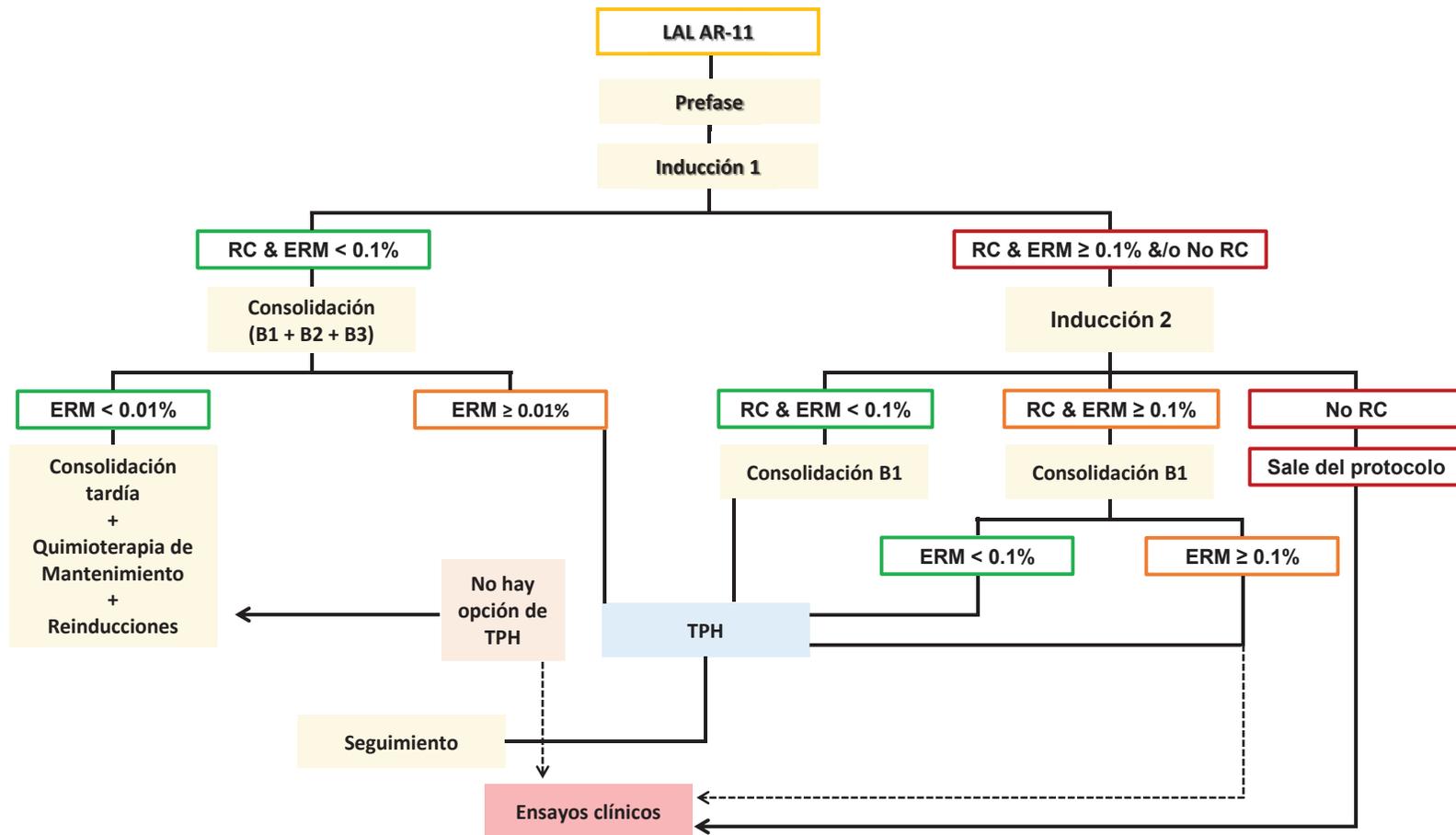
La supervivencia global (en inglés *overall survival*, OS) se definió como el tiempo entre la fecha de diagnóstico de la LAL y la fecha de último control, siendo toda muerte un evento. Se evaluó el impacto en la OS de las alteraciones, bien mutaciones o bien CNV, recurrentes en al menos 5 pacientes. Las curvas de la OS se representaron usando el método Kaplan-Meier y se compararon mediante el *test* log-rank. El análisis multivariante de factores pronósticos se realizó mediante el modelo de Cox de riesgos proporcionales.

La incidencia acumulada de recaída (en inglés *cumulative incidence of relapse*, CIR) se definió como el tiempo entre la fecha de la obtención de la RC y la fecha de recaída, considerando como evento competitivo toda muerte no relacionada con la recaída de la LAL. También se evaluó el impacto de las alteraciones recurrentes en la CIR. La CIR se estimó mediante funciones de incidencia acumulada y las curvas se compararon mediante el *test* de Gray.

Se consideró como significación estadística el valor de  $p < 0,05$ . Todos los análisis estadísticos se realizaron usando el software/paquete estadístico SPSS versión 24.0 (IBM Corp. Armonk, NY, USA) y el entorno de programación R versión 4.2.0. Estos cálculos fueron realizados por la estadística del grupo.



**Figura 22. Esquema del protocolo AR03.** En amarillo se indican los bloques de quimioterapia, en rojo los malos respondedores, en verde los buenos respondedores y en naranja los pacientes con ERM positiva. RC: remisión completa, ERM: enfermedad residual medible. TPH: Trasplante de progenitores hematopoyético.



**Figura 23. Esquema del protocolo AR11.** En amarillo se indican los bloques de quimioterapia, en rojo los malos respondedores, en verde los buenos respondedores y en naranja los pacientes con EMR positiva. RC: remisión completa, ERM: enfermedad residual medible. TPH: Trasplante de progenitores hematopoyético

## METODOLOGÍA DEL OBJETIVO 3

### **1. Pacientes y diseño del estudio**

En este objetivo se trabajó con muestras de 33 pacientes adultos con LLA-T tratados según 4 protocolos diseñados por el grupo PETHEMA: 2 pacientes se trataron según el protocolo LAL-07 FRAIL [NCT01366898], 2 pacientes según el protocolo AR-03 [NCT00853008], 21 pacientes según el protocolo AR-11 [NCT01540812] y 8 pacientes según el protocolo LAL19 [NCT04179929]. Para cada uno de los pacientes se analizó la muestra en el momento del diagnóstico (DX) y la muestra en el momento de la recaída (RE), teniendo en total 33 pares DX-RE (dobletes) (Tabla 8). Del total de pacientes, 21 estaban incluidos en la cohorte del Objetivo 2. Nuevamente, se obtuvo el consentimiento informado de todos los pacientes incluidos en el estudio. Las muestras y los datos clínicos se almacenaron de acuerdo con la declaración de Helsinki. Este estudio fue aprobado por la *Institutional Review Board* del Hospital Germans Trias i Pujol.

### **2. Detección y análisis de las alteraciones tipo SNV/indel**

Del total de los 33 dobletes, 15 pares DX-RE se secuenciaron usando el panel de genes diseñado, siguiendo los pasos explicados en el apartado 1.3 de la metodología del Objetivo 1. El resto de las muestras pareadas (n=18) se habían secuenciado previamente siguiendo una aproximación de tipo WGS. Las librerías se prepararon con el *kit KAPA HyperPrep* (Roche Kapa Biosystems, Basilea, Suiza) y la secuenciación se realizó en un *HiSeq 4000* o *NovaSeq 6000* (Illumina, California, EE. UU.) siguiendo un método *paired-end*, con un tamaño de lecturas de 2 x 151 pb. Ambos procedimientos se realizaron en el Centro Nacional de Análisis Genómico (CNAG, Barcelona)<sup>415</sup>.

El procesamiento de los datos y el filtrado de las variantes obtenidas por WGS se realizó siguiendo los mismos pasos que para el caso de la secuenciación tipo TDS, como se ha explicado en el apartado 1.3.4 de la metodología del Objetivo 1. Nuevamente las VAF de todas las variantes se corrigieron como se ha indicado en el apartado 1.5 de la metodología del Objetivo 1.

### **3. Análisis de las variaciones en el número de copias (CNV)**

Para analizar las alteraciones tipo CNV, se realizó *SNP-arrays* de 25 dobletes, siguiendo los pasos indicados en el apartado 1.4 de la metodología del Objetivo 1.

Para 6 de las muestras pareadas, la información de las alteraciones CNV se obtuvo a partir de los resultados de WGS utilizando la herramienta FACETS<sup>503</sup>. Este análisis lo realizó el técnico bioinformático del grupo. A partir de los datos crudos obtenidos (*raw data*), se validaron las alteraciones CNV obtenidas del análisis WGS. Para ello, se seleccionaron 10 muestras de las que se tenía información de las alteraciones CNV obtenidas por *SNP-arrays* y obtenidas por WGS, se compararon y se observó que ambas herramientas identificaban las mismas alteraciones.

Hubo 2 muestras pareadas de las que no se tuvo información de las alteraciones tipo CNV

**Tabla 8. Muestras pareadas DX-RE**

ID	Protocolo	DX					RE				
		Origen	% blastos	Técnica SNV/indel	Profundidad de lectura (X)	Técnica CNV	Origen	% blastos	Técnica SNV/indel	Profundidad de lectura (X)	Técnica CNV
PAT1	AR-11	MO	99	WGS	95	WGS	MO	78	WGS	95	WGS
PAT2	AR-11	SP	100*	TDS	335	SNP-arrays	MO	93,5	TDS	1190	SNP-arrays
PAT3	AR-11	MO	100	WGS	95	SNP-arrays	MO	100*	WGS	95	SNP-arrays
PAT6	AR-11	MO	100*	WGS	95	WGS	MO	100*	WGS	95	WGS
PAT7	AR-03	MO	98	TDS	640	WGS	MO	90	TDS	591	WGS
PAT8	AR-11	MO	98	WGS	95	SNP-arrays	MO	90	WGS	95	SNP-arrays
PAT9	AR-11	SP	100	WGS	95	WGS	SP	100*	WGS	95	WGS
PAT10	AR-11	MO	94	WGS	95	SNP-arrays	MO	36	WGS	95	SNP-arrays
PAT11	AR-11	MO	98	WGS	95	SNP-arrays	MO	98	WGS	95	SNP-arrays
PAT12	AR-11	MO	100*	WGS	95	SNP-arrays	MO	100*	WGS	95	SNP-arrays
PAT13	AR-11	SP	100*	WGS	95	SNP-arrays	MO	100*	WGS	95	SNP-arrays
PAT14	AR-11	MO	95	WGS	95	SNP-arrays	MO	95	WGS	95	SNP-arrays
PAT15	AR-11	MO	95	WGS	95	SNP-arrays	MO	100*	WGS	95	SNP-arrays
PAT16	AR-11	MO	83	WGS	95	WGS	MO	81	WGS	95	WGS
PAT18	AR-11	MO	73	WGS	95	WGS	MO	48	WGS	95	WGS
PAT19	FRAIL-07	MO	95	WGS	95	SNP-arrays	MO	80	WGS	95	SNP-arrays
PATE	LAL-19	MO	97	TDS	434	SNP-arrays	MO	57	TDS	1098	SNP-arrays
PATB	AR-11	SP	100	TDS	298	SNP-arrays	MO	75	TDS	665	SNP-arrays
PATO	AR-03	MO	55	TDS	166	SNP-arrays	MO	37	TDS	489	SNP-arrays
PATJ	AR-11	MO	95	TDS	538	SNP-arrays	MO	96	TDS	337	SNP-arrays
PATG	AR-11	MO	100	TDS	364	SNP-arrays	MO	100	TDS	440	SNP-arrays
PATI	LAL-19	MO	85.4	TDS	535	SNP-arrays	MO	71	TDS	451	SNP-arrays
PATH	AR-11	MO	68	TDS	273	SNP-arrays	MO	24	TDS	630	SNP-arrays
PATN	AR-11	SP	56	TDS	337	ND	MO	100	TDS	536	ND
PATM	LAL-19	MO	92	TDS	636	SNP-arrays	MO	100	TDS	439	SNP-arrays
PATD	LAL-19	SP	56	TDS	126	SNP-arrays	MO	100	TDS	205	SNP-arrays
PATL	LAL-19	MO	94	TDS	864	SNP-arrays	SP	81	TDS	942	SNP-arrays
PATP	AR-11	MO	99	TDS	354	SNP-arrays	MO	97	TDS	686	SNP-arrays
PATF	LAL-19	MO	94	TDS	864	SNP-arrays	MO	81	TDS	942	SNP-arrays
PATR**	FRAIL-07	MO	100	TDS	125	SNP-arrays	MO	88	TDS	752	SNP-arrays
PATQ	AR-11	MO	90	TDS	740	SNP-arrays	MO	37	TDS	523	SNP-arrays
PATA	LAL-19	SP	77	TDS	487	ND	MO	100	TDS	555	ND
PATK	LAL-19	MO	90	TDS	740	SNP-arrays	MO	97	TDS	525	SNP-arrays

\*Muestras sorteadas. \*\*La muestra de la recaída corresponde a una 2º recaída

#### 4. Análisis y seguimiento de mutaciones con PCR digital en muestras seriadas

El fundamento de la PCR digital (dPCR) consiste en repartir la muestra de ADN en miles de pocillos, de tal manera que solo caiga una o dos moléculas en cada pocillo, donde ocurre la reacción de PCR de forma individual. Para los ensayos realizados en esta tesis se utilizaron sondas con química *TaqMan* que son específicas de una secuencia diana y están marcadas con un fluorocromo. Las sondas que se unen al alelo no mutado (en inglés *wild type*, WT) están marcadas con el fluorocromo VIC, mientras que las sondas que se unen al alelo mutado están marcadas con el fluorocromo FAM. De este modo, en función del color de la señal emitida se conoce si la molécula de ADN que está presente en un determinado pocillo presenta o no la mutación.

En este trabajo, se utilizó la tecnología de dPCR basada en chips de nanofluídica *QuantStudio™ 3D Digital PCR 20K Chip Kit v2* (cat. no. A26316, Thermo Fisher Scientific, Massachusetts, EE. UU.). Cada chip contiene 20.000 pocillos donde se dan 20.000 reacciones independientes de PCR.

Las mutaciones a detectar fueron: *NRAS\_Q61K* y *NT5C2\_R367Q*. Para la mutación de *NRAS* había disponible una sonda comercial (cat. no. A44177 Thermo Fisher Scientific, Massachusetts, EE. UU.) y para la mutación de *NT5C2* se diseñó un ensayo mediante la herramienta *Custom TaqMan® Assay Design Tool* (Thermo Fisher Scientific, Massachusetts, EE. UU.).

Dado que el objetivo era alcanzar la mayor sensibilidad posible, se realizaron pruebas para determinar la mayor cantidad de ADN que se podía incluir en cada chip sin llegar a la concentración de saturación. La cantidad de ADN determinada fue 300 ng cuantificado por *Nanodrop*.

Se preparó la reacción de PCR, donde además del ADN se incluyó la *Master Mix Quant St 3D MM 2X* (cat. no. A26358, Thermo Fisher Scientific, Massachusetts, EE. UU.), las sondas específicas del ensayo 20X y agua libre de nucleasas. En todas las determinaciones se añadió un control negativo (reacción sin ADN).

En cada chip se cargaron 14,5 µl de reacción con la ayuda del *QuantStudio™ 3D Digital PCR Chip Loader* (cat. no. 4482592 Thermo Fisher Scientific, Massachusetts, EE. UU.). La PCR se realizó en el termociclador *GeneAmp 9700* (cat. no. 4339386, Thermo Fisher Scientific, Massachusetts, EE. UU.). En el caso del ensayo *NT5C2* se probaron diferentes condiciones de ciclos y temperaturas, pero los mejores resultados se obtuvieron siguiendo las recomendaciones del fabricante. En el caso del ensayo *NRAS* también se siguieron las recomendaciones del fabricante. Una vez finalizado el programa de termociclador, se leyó la fluorescencia de los chips con el lector *QuantStudio 3D dPCR Instrument* (cat. no. 4489084, Thermo Fisher Scientific, Massachusetts, EE. UU.).

Los resultados se analizaron usando el programa *QuantStudio 3D AnalysisSuite Cloude* (Thermo Fisher Scientific, Massachusetts, EE. UU.).

Previamente a analizar las muestras de interés, se determinaron dos parámetros necesarios para poner a punto el ensayo: límite de blanco (en inglés *limit of blank*, LOB) y límite de detección (en inglés *limit of detection*, LOD).

El LOB determina el número de pocillos positivos que se detectan en una muestra WT para la mutación estudiada (falsos positivos). A partir de este se establece el mínimo número de pocillos que se requieren para tener unos resultados fiables. Para determinar este valor se analizaron un total de 8 muestras WT para cada una de las mutaciones. En el caso del ensayo de *NRAS* el LOB fue de 2 pocillos y para el caso del ensayo de *NT5C2* fue de 6 pocillos (Tabla 9).

El LOD determina el menor número de pocillos positivos que se pueden detectar en un ensayo, es decir, la sensibilidad de la técnica. Para determinar el LOD, para cada ensayo, se hicieron diluciones seriadas de un control positivo para la mutación con una muestra WT. La VAF del control positivo se conocía porque se había secuenciado previamente. En el caso del ensayo *NRAS* el LOD fue 0,004% y para el ensayo *NT5C2* fue 0,01% (Tabla 10).

En total, se analizaron 18 muestras pertenecientes a 4 pacientes (Tabla 11). Los resultados de la dPCR se expresaron como porcentaje de copias con la mutación respecto al total de copias de ADN.

**Tabla 9. Determinación LOB dPCR**

Ensayo	Muestra	Nº chips analizados	Copias VIC	Copias / µl (VIC)	Copias FAM	Copias/ µl (FAM)	Mínimos puntos positivos	Media de Falsos positivos	Desviación estándar (Falso positivos)
<i>NRAS</i> Q61K	WT <sub>1</sub>	1	135447,4	9341,2	1,1	7,57x10 <sup>-2</sup>	2	0,57	1
	WT <sub>2</sub>	1	25776,65	1777,7	1,1	7,41x10 <sup>-2</sup>			
	WT <sub>3</sub>	2	88306,45	6090,1	0,0	0			
	WT <sub>4</sub>	2	59112,15	4076,7	2,5	0,172			
	WT <sub>5</sub>	2	3340,8	230,4	1,1	7,72x10 <sup>-2</sup>			
	WT <sub>6</sub>	1	82702,2	5703,6	0,0	0			
	WT <sub>7</sub>	2	92676,75	6391,5	2,1	0,148			
	WT <sub>8</sub>	1	90881,65	6267,7	0,0	0			
<i>NT5C2</i> R367Q	WT <sub>1</sub>	1	71885,2	4957,6	3,2	0,2	6	2,7	1,7
	WT <sub>2</sub>	2	82418	5684	2,4	0,2			
	WT <sub>3</sub>	2	86325,75	5953,5	4,0	0,3			
	WT <sub>4</sub>	1	89934,8	6202,4	2,2	0,1			
	WT <sub>5</sub>	1	68954,75	4755,5	3,3	0,2			
	WT <sub>6</sub>	2	59384,75	4095,5	4,2	0,3			
	WT <sub>7</sub>	1	53126,55	3663,9	2,1	0,1			
	WT <sub>8</sub>	1	59051,25	4072,5	5,2	0,4			

**Tabla 10. Determinación LOD dPCR**

Ensayo	LOB establecido	Diluciones	Nº chips analizados	Copias VIC	Copias/µl (VIC)	Copias FAM	Copias/µl (FAM)	Mutación/ Total (%)	Mutación esperada/ total (%)
NRAS Q61K	2	1:1	1	33305,05	2296,9	40858,1	2817,8	55	50
		1:5	1	50776,1	3501,8	7218,4	497,8	12	10
		1:50	1	59864,7	4128,6	701176,5	48357,0	1	1
		1:500	2	53307,8	3676,4	68063,0	4694,0	0,1	0,1
		1:5.000	3	57534,55	3967,9	7,8	0,5	0,01	0,01
		1:7.000	3	56313,65	3883,7	1,9	0,1	0,003	0,007
		1:10.000	<b>3</b>	<b>59829,9</b>	<b>4126,2</b>	<b>3,0</b>	<b>0,2</b>	<b>0,004</b>	<b>0,005</b>
		1:17.000	3	60508,5	4173	1,1	0,1	0,002	0,003
NT5C2 R367Q	6	1:1	1	43733,5	3016,1	36064,4	2487,2	45	42
		1:4	2	70122,0	4836	8290,5	571,76	11	10
		1:40	4	55342,2	3816,7	56999,5	3.931	0,1	0,1
		1:400	<b>6</b>	<b>58387,2</b>	<b>4026,7</b>	<b>7,7</b>	<b>0,533</b>	<b>0,01</b>	<b>0,01</b>
		1:5.000	6	57643,3	3975,4	4,8	0,332	0,008	0,007
		1:8.000	6	58394,4	4027,2	5,2	0,356	0,007	0,005

**Tabla 11. Muestras analizadas dPCR**

Paciente	Punto de tratamiento	Días desde el DX	Origen	ERM (%)	Mutación	Nº de chips analizados
PAT 18	Inducción +15	15	MO	0,006	NRAS Q61K	6
	Final inducción	37	MO	0,001	NRAS Q61K	6
	Tras Cons temprana 1	66	MO	0,003	NRAS Q61K	2
	Tras Cons temprana 3	123	MO	0,001	NRAS Q61K	3
	Recaída	148	MO	46	NRAS Q61K	2
PAT F	Diagnóstico	0	MO	94	NT5C2 R367Q	3
	Inducción +15	22	MO	0,1	NT5C2 R367Q	2
	Final inducción	49	MO	0,01	NT5C2 R367Q	4
	Tras Cons temprana 3	155	MO	0,0002	NT5C2 R367Q	3
	Tras Cons tardía 3	301	MO	0,0002	NT5C2 R367Q	6
	Recaída	648	MO	42	NT5C2 R367Q	1
PAT 14	Diagnóstico	0	MO	95	NT5C2 R367Q	2
	Tras Cons temprana 3	110	MO	<0,01	NT5C2 R367Q	3
	Tras Cons tardía 3	199	MO	<0,001	NT5C2 R367Q	4
	Recaída	340	MO	95	NT5C2 R367Q	1
PAT 16	Tras Cons temprana 3	146	MO	0,001	NT5C2 R367Q	6
	Tras Cons tardía 3	237	MO	0,001	NT5C2 R367Q	6
	Recaída	292	MO	81	NT5C2 R367Q	2

DX: Diagnóstico; ERM: Enfermedad mínima residual cuantificada por citometría de flujo multiparamétrica (CFM); Cons: consolidación, MO: Médula ósea

## 5. Amplificación de leucemias LLA-T primarias en ratones inmunodeprimidos (PDX)

Con el objetivo de generar un banco de células leucémicas LLA-T se inyectaron 30 LLA-T primarias en un total de 70 ratones de la cepa NSG (46 hembras y 24 machos, de 8-10 semanas de edad, *NOD/SCID/IL-2R null*).

Estos ratones se habían irradiado previamente a una dosis subletal de 2,5 grays (Gy). El número de células inyectadas fue entre  $5000-10^6$  y se inyectaron bien en el seno retro-orbital o intra-tibial. Previamente a la inyección (30 minutos) se inyectó en cada ratón una combinación de buprenorfina (0,05-0,1 mg/kg) y meloxicam (2 mg/kg), vía subcutánea.

Posteriormente, se anestesiaron con isoflurano (inducción: 4,5-5% isoflurano + 1-1,5 L oxígeno/minuto; mantenimiento del estado anestésico: 2% isoflurano + 0,8-1 L/min oxígeno). En el caso de la administración vía intra-tibial, se inyectaron 30  $\mu$ l de células leucémicas en PBS y 10% FBS usando una aguja de 26G. En el caso de la administración retro-orbital se inyectaron las células leucémicas usando una aguja de 26G-29G. En la tabla 12 se muestra el número de células inyectadas para los ratones donde se inyectó correctamente la leucemia.

Tras 6 semanas después de la infusión de las células leucémicas, se comenzó a evaluar periódicamente el desarrollo de la leucemia en el ratón. Con este fin, cada 15 días se extrajo sangre (200  $\mu$ l aproximadamente) mediante una punción maxilofacial y se recolectó en un tubo EDTA. Se añadió el mismo volumen de PBS 1X que de sangre y se centrifugó durante 10 minutos a 1000 x g. A continuación, se retiró el sobrenadante y se añadió 1 mL de tampón ACK (amoníaco-cloruro-potasio) (cat. no. 10-548E Lonza Bioscience, Maryland, EE. UU.) para lisar los eritrocitos. Este tampón se mantuvo durante 30 segundos en la muestra y finalmente se añadió 10 mL de PBS 1X para inactivarlo. Las muestras se centrifugaron durante 10 minutos a 300 x g y se resuspendieron en 200  $\mu$ l de PBS 1X + 2% FBS. Se contaron las células usando una cámara Neubauer.

Posteriormente, se marcaron con los anticuerpos anti-CD45 (APC Mouse Anti-Human CD45) (cat. no. 555485), anti-CD7 (PE Mouse Anti-Human CD7) (cat. no. 555361) y anti-HLA-ABC (FITC Mouse Anti-Human) (cat. no. 555552) (BD Biosciences), añadiendo 1:5  $\mu$ l de anticuerpo por cada millón de células. La incubación con los anticuerpos se realizó durante 30 minutos y a continuación, se hizo un lavado con PBS 1X, resuspendiendo las células en volumen final de 200  $\mu$ l de PBS 1X que se analizaron en el citómetro *FACSCanto II* (BD Biosciences)). El programa utilizado para el análisis de las poblaciones celulares fue FlowJo\_V10.

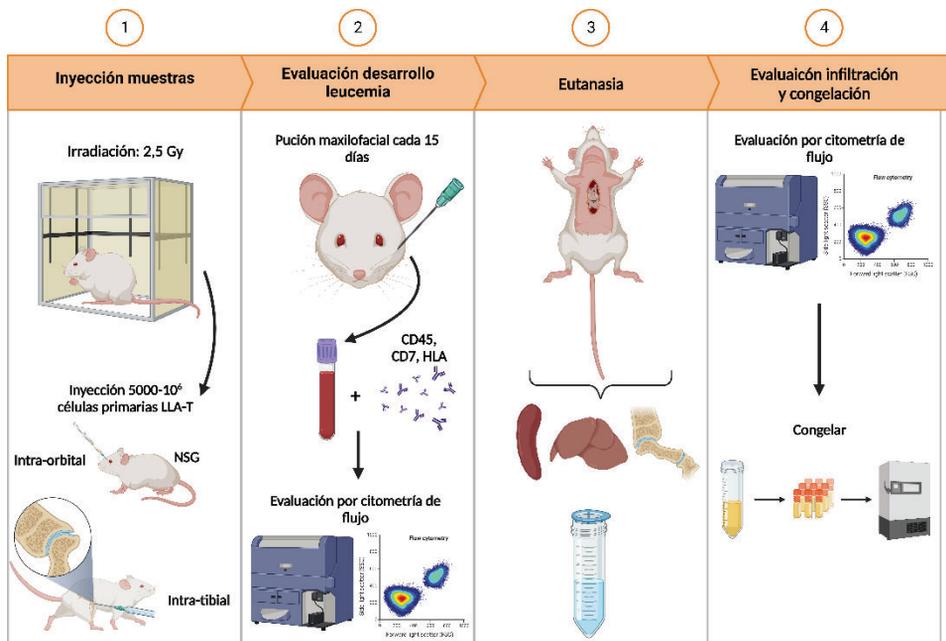
Los ratones se sacrificaron, mediante dislocación cervical, cuando presentaron un porcentaje de blastos >20% en sangre o mostraron signos claros de enfermedad. Esto ocurrió en un intervalo entre 9-30 semanas.

Se aisló la médula ósea, bazo e hígado y en algunos casos ganglios linfáticos. Se partió del fémur y la tibia para extraer la médula inyectando PBS 1X con una jeringa en el interior del hueso; la suspensión celular obtenida se pasó por un filtro de 45  $\mu$ m.

En el caso del brazo, hígado y ganglio linfático se disgregaron los órganos mediante el embolo de una jeringa de 20 ml y se filtraron añadiendo PBS 1X con un filtro de 45 µm. La solución obtenida se contó con la cámara Neubauer.

Posteriormente, las muestras se evaluaron por citometría de flujo para confirmar la presencia de infiltración con los marcadores hCD7, hCD45 y hHLA siguiendo el procedimiento explicado anteriormente.

En caso de infiltración todo el material obtenido se congeló a una concentración de 10 millones/mL en un medio que incluyó 50% de RPMI (cat. no.21870076) (Thermo Fisher Scientifics, Massachusetts, EE. UU.) + 10% DMSO y 50% de FBS (cat. no. 10270, Thermo Fisher, Massachusetts, EE. UU.). En la figura 24 se muestra el diagrama de este proceso.



**Figura 24. Amplificación de leucemias LLA-T en ratones NSG**

Del total de leucemias inyectadas se consiguieron amplificar 10 (33,3%) con resultados óptimos (infiltración >40% blastos) y 6 (18%) con un porcentaje de infiltración menor (<10%). En la tabla 12 se muestra un resumen del material inyectado en cada ratón y los datos de infiltración en cada órgano para las leucemias amplificadas óptimamente.

En esta tesis se utilizó el material de los PDX obtenidos a partir de las muestras PAT2, PAT6, PAT9, PAT15.

**Tabla 12. Muestras expandidas en ratones NSG**

ID	N <sup>o</sup> células inyectadas/ Vía de administración	Origen de la muestra inyectada	Punto final (Semanas desde la inyección)	% infiltración			
				MO	Bazo	Hígado	Ganglio
PAT4	90.000 (IT)	MO	30	70	10	10	4
PAT6	15.000 (IT)	MO	16 16	82	80	85	-
	15.000 (IT)	MO		80	85	85	-
PAT9	400.000(IT)	MO	21	2	-	5	93
PAT2	400.000(RET)	SP	21	92	65	77	-
	400.000(IT)	SP	22	90	56	7	-
PAT25	1,3X10 <sup>6</sup> (RET)	-	19	95	41	11	-
	1,2X10 <sup>6</sup> (RET)	-	19	90	69	13	-
	400.000(IT)	-	18	30	37	10	-
	400.000(IT)	-	17	89	50	20	-
PAT27	1,3X10 <sup>6</sup> (RET)	MO	8	73	77	40	-
PAT15	1,3X10 <sup>6</sup> (RET)	MO	9	93	65	75	-
	1,3X10 <sup>6</sup> (RET)	MO	11	80	75	65	-
PAT20	2X10 <sup>6</sup> (RET)	MO	28	70	-	-	-
PAT21	200.000(RET)	MO	27	40	-	15	-
PAT22	1X10 <sup>6</sup> (RET)	MO	20	40	-	-	-

RET: retro-orbital; IT: intra-tibial; MO: médula ósea; SP: sangre periférica

## 6. Análisis estadísticos

Los análisis estadísticos entre las muestras pareadas DX-RE se realizaron mediante el *test* no paramétrico de Wilcoxon. El análisis de la relación entre variables no paramétricas e independientes se realizó mediante el *test* de Mann-Whitney.

La comparación entre las medianas de muestras no paramétricas se hizo mediante el *test* comparación de las medianas. Las correlaciones entre variables no paramétricas se realizaron con el *test* de Spearman.

Para estudiar las asociaciones positivas o negativas entre las variables se calculó el *odds ratio*, que se representó como  $\log_{10}(\text{odds ratio})$ . En aquellos casos en los que el denominador del *odds ratio* resultó ser 0, se hizo una corrección sumando 0,5 a cada término para calcular el  $\log_{10}(\text{odds ratio})$ .

Se consideró como significación estadística  $p < 0,05$ . Se utilizó el software GraphPad Prism 8 para todos los análisis y para la elaboración de gráficos y figuras.





## RESULTADOS



## RESULTADOS DEL OBJETIVO 1

### 1. Caracterización genómica de los pacientes con LLA-T en el momento del diagnóstico.

Con el fin de identificar alteraciones implicadas en la patogénesis de la LLA-T de pacientes adultos se analizaron mediante TDS un total de 146 pacientes con LLA-T para detectar SNV e indels de < 50 pb. Además, para 134 pacientes también se analizaron las alteraciones CNV mediante SNP-*arrays*. Del total de pacientes analizados por TDS, 9 no presentaron ninguna alteración, pero de éstos hubo 6 que presentaron CNV. De los 134 pacientes analizados por SNP-*arrays*, 6 no presentaron ninguna CNV, pero 5 sí que presentaron SNV o indels.

Así pues, teniendo en cuenta los datos de SNV, indels y CNV, 3 pacientes no presentaron ninguna de estas alteraciones, aunque para 2 de ellos no se tenía la información del análisis de CNV (Fig. 25).

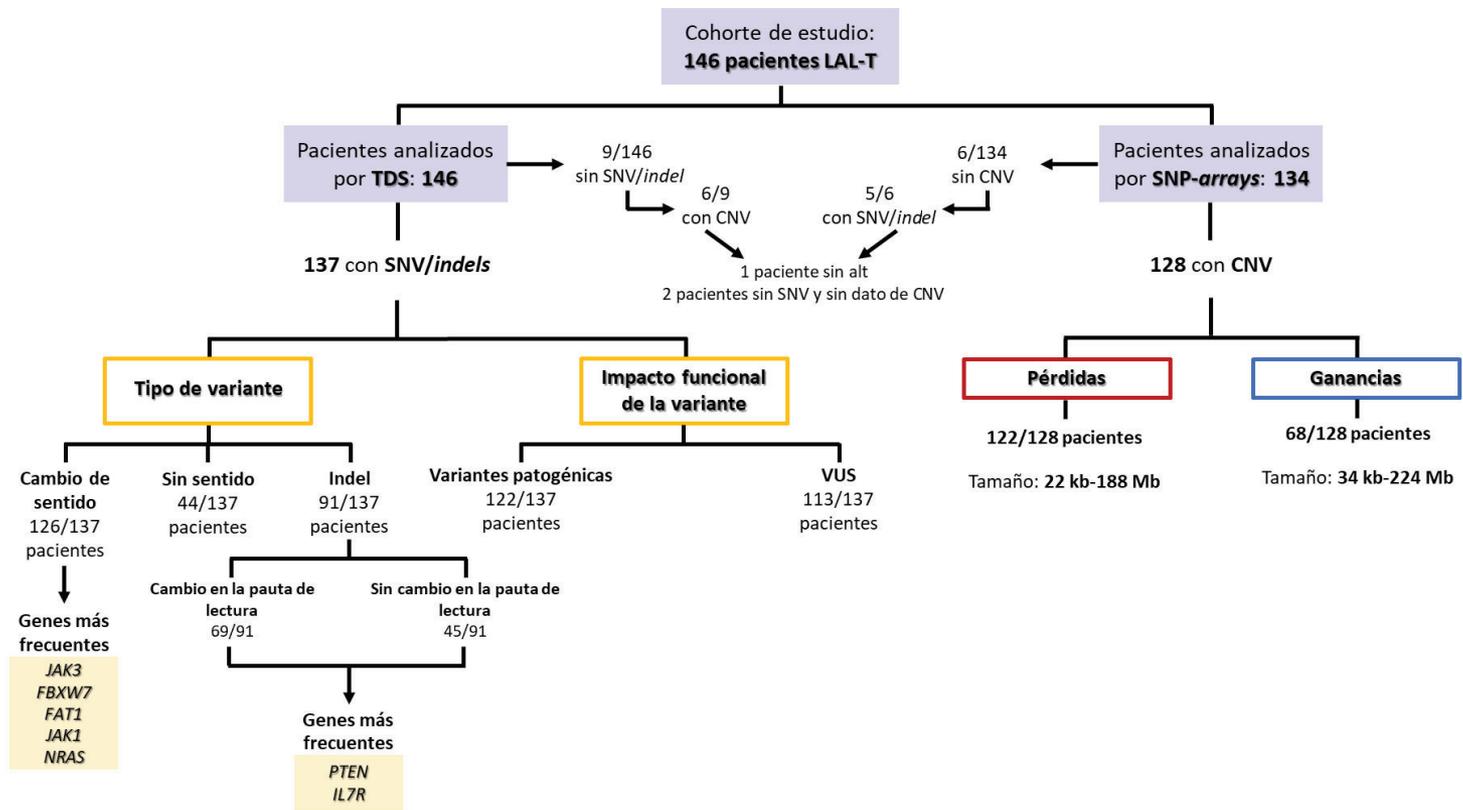
#### 1.1. Alteraciones detectadas por TDS

En el caso de la secuenciación, la profundidad media de lectura fue de 286X. Para los 137 pacientes con alteraciones se identificaron un total de 533 variantes, presentando una mediana [extremos] de 4 [1-12] variantes/paciente y 3 [1-8] genes mutados/paciente.

Los genes *NOTCH1*, *PTEN*, *IL7R*, *JAK3*, *JAK1* y *DNMT3A* mostraron más de una mutación de media por paciente. Un 92% (126/137) de los pacientes presentaron variantes con cambio de sentido (*missense*), un 32% (44/137) variantes sin sentido (*nonsense*) y un 66% (91/137) variantes tipo indel; dentro de este tipo de variantes el 76% (69/91) de los pacientes presentaron indels que afectaron a la pauta de lectura (*frameshift*) y el 49% (45/91) presentaron indels que no afectaron a la pauta de lectura (*non-frameshift*).

Globalmente, el 62% (85/137) de los pacientes presentaron la combinación de SNV e indels, mientras que un 34% (46/137) presentaron solo SNV, y un 4% (6/137) de los pacientes presentaron únicamente variantes de tipo indels (3 pacientes de tipo *frameshift* y 3 de tipo *non-frameshift*). Hubo genes que presentaron principalmente variantes tipo *missense*, fue el caso de *JAK3* (96,3%), *FBXW7* (92,3%), *FAT1* (100%), *JAK1* (100%) y *NRAS* (100%); mientras que otros genes presentaron más comúnmente variantes tipo *indel*, fue el caso de *PTEN* (71,4%) e *IL7R* (80%).

De acuerdo con la clasificación de las variantes según su posible impacto funcional, un 89% (122/137) de los pacientes presentaron variantes anotadas como patogénicas mientras que el 80% (113/137) presentaron variantes VUS. Solo 15 pacientes presentaron exclusivamente variantes VUS. La práctica totalidad de variantes anotadas como patogénicas fueron *missense* (98%), mientras que las variantes VUS la mayoría fueron indels (62%), pues para este tipo de variantes los predictores generalmente no dan información sobre su posible repercusión; el resto de las variantes de significado incierto fueron 19% *missense* y 19% *nonsense* (Fig 25).



**Figura 25. Diagrama del tipo y número de alteraciones en la cohorte de estudio (n=145).** TDS: *Target Deep Sequencing* (Secuenciación dirigida a alta profundidad); SNP: *SNP-arrays* (matrices de SNPs); SNV: *Single Nucleotide Variant* (Variante de un único nucleótido); CNV: *Copy Number Variation* (Variación en el número de copias).

## 1.2. Alteraciones detectadas por SNP-arrays y determinación de los genes diana en las alteraciones tipo CNV

En el análisis de las CNV, considerando los 128 pacientes que presentaron alteraciones, se identificaron un total de 674 alteraciones de las cuales un 79% (534/674) fueron pérdidas y un 21% (140/674) ganancias. La mediana [extremos] de alteraciones por pacientes fue de 3 [1-20] para el caso de las pérdidas y 1 [1-15] para el caso de las ganancias. Un 95% (122/128) de los pacientes presentaron al menos una pérdida y un 53% (68/128) de los pacientes presentaron al menos una ganancia. La pérdida más pequeña fue de 22 kb y se correspondió con la delección focal del gen *LEF1* y la más grande tuvo 188 Mb de tamaño y se correspondió con una monosomía del cromosoma 4. La ganancia más pequeña tuvo un tamaño de 34 kb y se trató de la duplicación focal del gen *STAT5B*, la de mayor tamaño fue de 224 Mb y se correspondió con una trisomía del cromosoma 2.

Con el objetivo de identificar las CNV involucradas en el desarrollo de la LLA-T y posteriormente evaluar su impacto pronóstico (Objetivo 2) se seleccionaron aquellas alteraciones que afectaron al menos a 5 pacientes. Hubo alteraciones para las cuales las coordenadas de inicio y fin en los diferentes pacientes no fueron exactamente las mismas, y en muchas ocasiones tampoco los genes incluidos. En estos casos, se intentó identificar, en las que fue posible, el gen diana de la alteración. De este modo, se identificaron diferentes tipos de alteraciones con diferente tamaño y complejidad:

- **Aneuploidías:**

Fueron recurrentes las trisomías del cromosoma 10, trisomías del cromosoma 19 y las trisomías del cromosoma X (Tabla 13).

**Tabla 13. Aneuploidías**

Alteración	ID paciente
Trisomía Cromosoma 10	TALL31
	TALL55
	TALL99
	TALL77
	TALL113
	TALL6
	TALL52
Trisomía Cromosoma 19	TALL38
	TALL111
	TALL31
	TALL69
	TALL107
	TALL145
Trisomía Cromosoma X	TALL86
	TALL42
	TALL130
	TALL125
	TALL99
TALL69	

- **Alteraciones de brazos de cromosomas:**

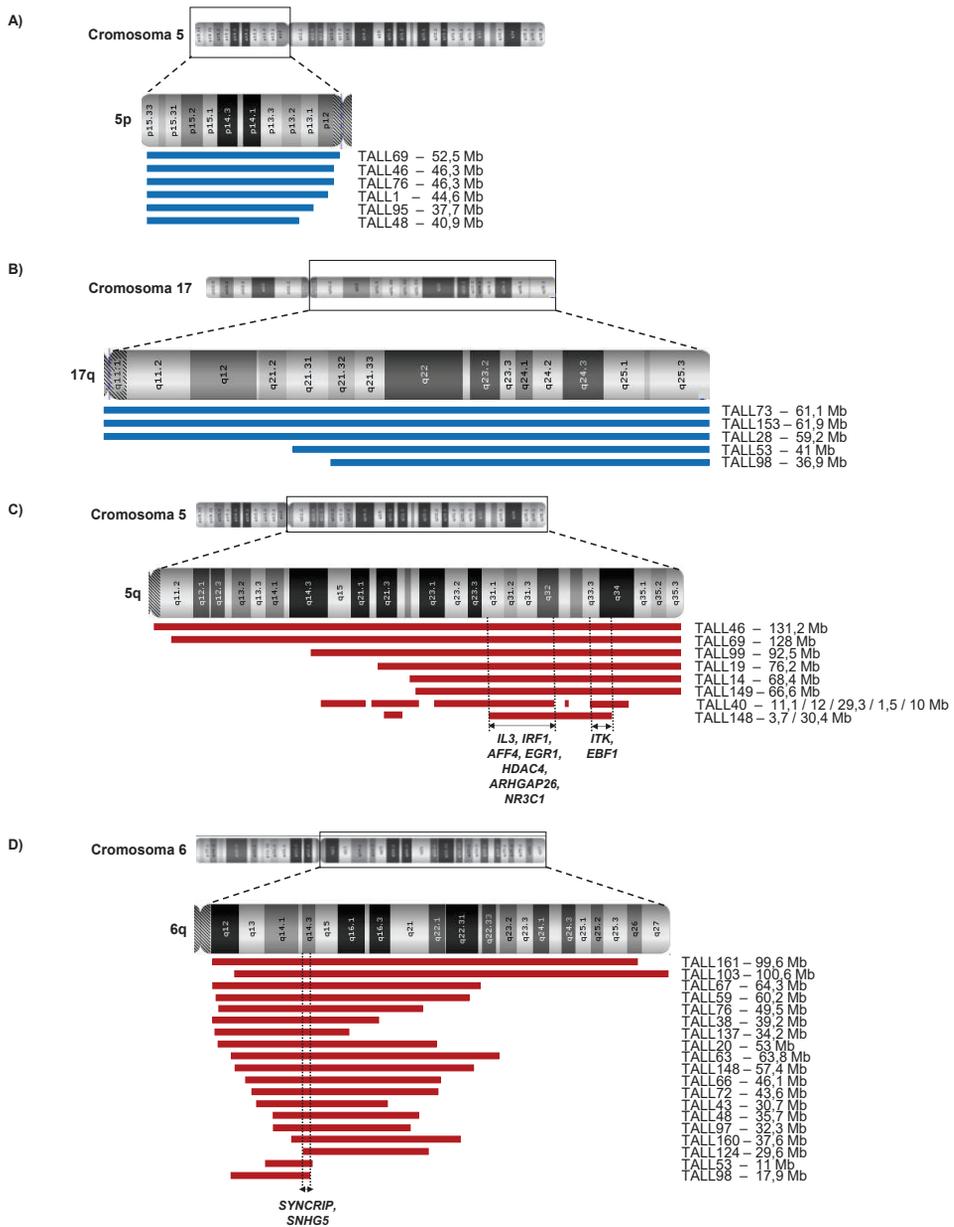
Se encontraron duplicados los brazos 5p (dup(5p)) (Fig. 26A) y 17q (dup(17q)) (Fig. 26B). Además, también fueron recurrentes las deleciones en la región del brazo 5q (del(5q)) (Fig. 26C) y 6q (del(6q)) (Fig. 26D). En ambos casos las coordenadas de inicio y fin para los diferentes pacientes afectados no fueron idénticas, pero se consideraron la misma alteración por presentar regiones altamente solapantes.

En el caso de las del(5q) se observaron dos regiones mínimamente solapantes que incluyeron genes importantes en cáncer y en la LLA-T. La primera incluyó los genes *IL3*, *IRF1*, *AFF4*, *EGR1*, *HDAC4*, *ARHGAP26* y *NR3C1*; mientras que la segunda los genes *ITK*, *EBF1* y *PTTG1* (Fig. 26C). En las del(6q) la mínima región solapante coincidió con la región reportada por Gachet S., El-Chaar T., Avran D., Genescà E. *et al.*<sup>360</sup>, incluyendo los genes *SYNCRIP* y *SNHG5* (Fig. 26D).

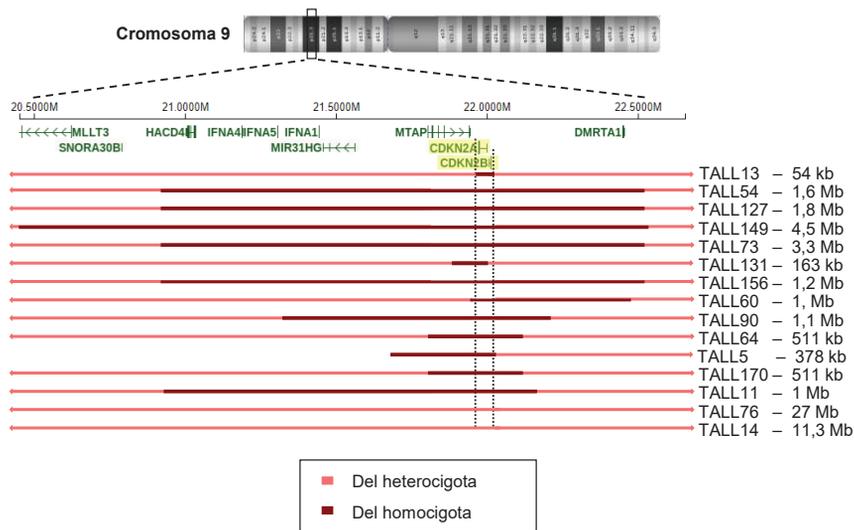
- **Alteraciones asociadas al inmunofenotipo**

Hubo un conjunto de alteraciones que, a pesar de tener coordenadas diferentes, la mínima región alterada común a todas ellas coincidió con un gen *driver* de la LLA-T, sugiriendo que este sería el gen diana de la alteración. No obstante, en muchos casos estas regiones incluyeron más de un gen *driver* lo que dificultó la identificación del gen diana. Fue el caso de las alteraciones encontradas en el brazo 9p, brazo 12p, brazo 13q y brazo 16p. Sin embargo, para todas ellas se observó que, del total de genes incluidos, había una asociación positiva entre uno de ellos y el inmunofenotipo de los pacientes que la presentaron, sugiriendo que dicho gen sería el responsable de desencadenar el bloqueo en la diferenciación del linfocito T y, por tanto, sería el gen diana de la alteración.

- En el brazo 9p, en la citobanda 9p21.3, se localizan los genes *CDKN2A/B*. La deleción de *CDKN2A/B* es la alteración más frecuente en la LLA-T<sup>504</sup>. En nuestra cohorte, de los 134 pacientes con resultados para CNV, 81 (60%) presentaron deleciones que afectaron a *CDKN2A/B*. El tamaño de estas deleciones fue variable, desde la deleción de la totalidad del brazo 9p (tamaño 67,8 Mb) hasta la deleción que solo incluía el gen *CDKN2B* (26,9 Kb). No obstante, solo 15 pacientes presentaron deleciones mayores a 5 Mb y en todos estos casos, excepto 2 (TALL76 y TALL14), se observó que la región que incluía los genes *CDKN2A/B* estaba delecionada en homocigosis, mientras que el resto de la región alterada lo estaba en heterocigosis (Fig. 27), sugiriendo que este gen sería la diana de la alteración. Del total de los 81 pacientes con deleciones en esta región, 72 (88,9%) tuvieron delecionados ambos genes, en 8 (10%) la deleción afectó únicamente a *CDKN2A* y solo un paciente presentó deleción únicamente de *CDKN2B*. La mayoría de los pacientes (69/81; 85%) presentaron deleción homocigota de *CDKN2A/B*. Teniendo en cuenta todas las deleciones, se observó que esta alteración era recurrente en los pacientes con inmunofenotipo cortical (OR deleción *CDKN2A/B*-inmunofenotipo cortical: 4,4; p= 0,0004), asociación descrita en la LLA-T<sup>112</sup>. Estos resultados sugirieron que, a pesar del tamaño variable de la región alterada, la deleción de los genes *CDKN2A/B* produjo un bloqueo del proceso de diferenciación en la etapa cortical de los linfocitos T, pudiéndose considerar estos como los genes diana de este conjunto de alteraciones.

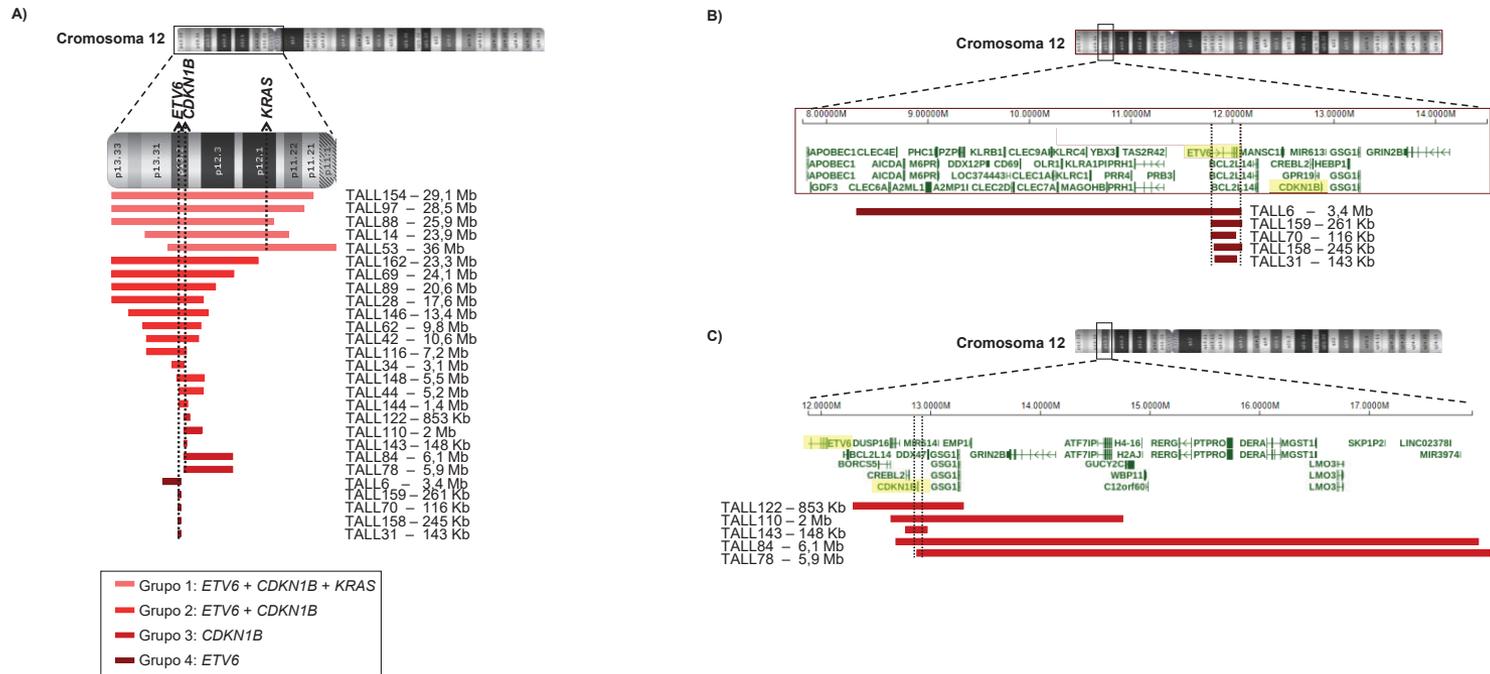


**Figura 26. Alteraciones de brazos de cromosoma.** A) Tamaño y posición de las duplicaciones que afectan al brazo 5q. B) Tamaño y posición de las duplicaciones que afectan al brazo 17q. C) Tamaño y posición de las deleciones que afectan al brazo 5q. D) Tamaño y posición de las deleciones que afectan al brazo 6q. Para cada alteración se indica el identificador del paciente (TALL) y el tamaño de la alteración. Para las del(5q) y del(6q) en la parte inferior se indican los genes comunes a las mínimas regiones solapantes.



**Figura 27. Deleciones de los genes *CDKN2A/B*.** Tamaño y posición de las deleciones que incluyen los genes *CDKN2A/B* que tuvieron un tamaño mayor a 5 Mb. Las regiones con deleción homocigota se representan en granate. Los genes *CDKN2A* y *CDKN2B* se indican en amarillo. Para cada alteración se indica el identificador del paciente (TALL).

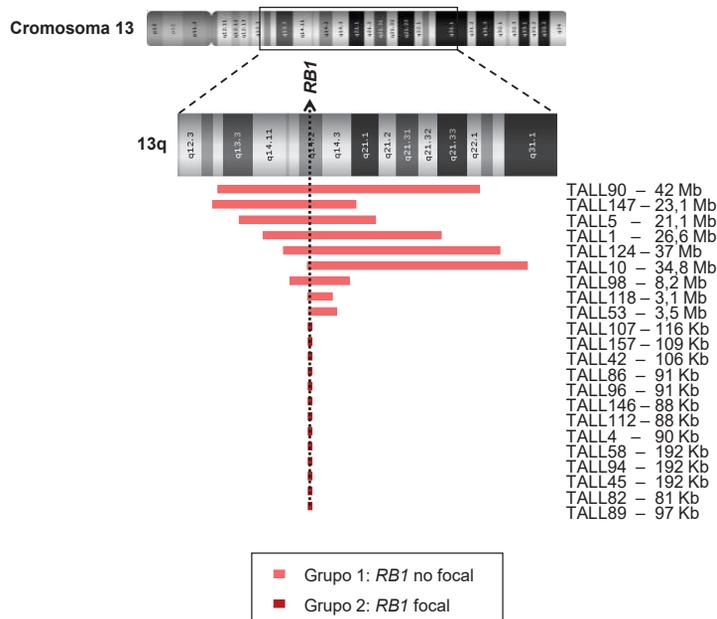
- El brazo 12p fue una región compleja de analizar ya que contiene varios genes *drivers* para la LLA-T. En concreto, en la citobanda 12p13.2 se localizan *ETV6* y *CDKN1B* que se encuentran separados por 800 kb. Un total de 27 pacientes presentaron alteraciones en esta región, de tamaños variables pero solapantes (Fig. 28A). De estos, 5 pacientes presentaron deleciones focales de *ETV6* (Fig. 28A grupo 4 y Fig. 28B), 5 pacientes presentaron deleciones que solo afectaron a *CDKN1B* (Fig. 28A grupo 3 y Fig. 28C), 12 presentaron deleciones que incluyeron *ETV6* y *CDKN1B* (Fig. 28A, grupo 2) y 5 pacientes presentaron deleciones que incluyeron *ETV6*, *CDKN1B* y *KRAS* (Fig. 28A, grupo 1). Las deleciones focales de *CDKN1B* (n=5) se encontraron en pacientes con un inmunofenotipo cortical (OR deleción *CDKN1B*-inmunofenotipo cortical: 8,3; p=0,0053), asociación descrita anteriormente<sup>112</sup>. Este inmunofenotipo no es común en los pacientes con *ETV6* alterado<sup>197</sup> y, de hecho, ninguno de los pacientes con deleción focal de *ETV6* tuvo un inmunofenotipo cortical. En cambio, 8/12 pacientes que presentaron deleción tanto *ETV6* como *CDKN1B* tuvieron un inmunofenotipo cortical (p=0,018). Finalmente, el grupo de pacientes con deleciones que incluyeron *ETV6*, *CDKN1B* y *KRAS* no presentaron una asociación con un inmunofenotipo específico, sugiriendo que estas deleciones tenían una repercusión diferente. Por tanto, con esto se decidió agrupar las deleciones del brazo 12p en tres grupos: i) grupo *ETV6* focal (Fig. 28A grupo 4); ii) grupo *CDKN1B* (que incluyó deleciones focales de *CDKN1B* y deleciones que afectaron a *ETV6* y *CDKN1B*) (Fig. 28A grupo 2 y 3); iii) grupo con las deleciones que afectaron a *ETV6*, *CDKN1B* y *KRAS* (Fig. 28A grupo 1).



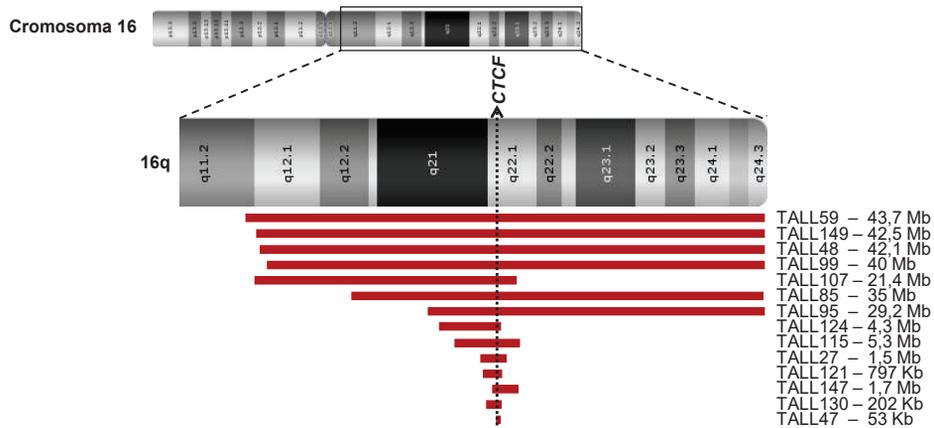
**Figura 28. Deleciones en el brazo 12p.** A) Tamaño y posición de las deleciones que afectan al brazo 12p. Cada subgrupo, definido de acuerdo con los genes que incluye la deleción, se representa en un color diferente. B) Tamaño y posición de las deleciones focales del gen *ETV6*. C) Tamaño y posición de las deleciones que incluyen *CDKN1B* y no *ETV6*. En amarillo (B y C) se marcan los genes *ETV6* y *CDKN1B*. Para cada alteración se indica el identificador del paciente (TALL).

- Las deleciones en el brazo 13q, donde se encuentra el gen *RB1* entre otros, fueron otras de las alteraciones recurrentes. En concreto, 22 pacientes presentaron deleciones de esta región. En 13 pacientes la deleción tuvo un tamaño pequeño (~ 200 kb) que únicamente incluyó al gen *RB1*, mientras que en 9 pacientes la deleción tuvo un tamaño mayor y, además del *RB1*, también incluyó otros genes (Fig. 29). Estudiando el inmunofenotipo de los pacientes con deleciones en el brazo 13q, se observó que 10/13 pacientes con deleción focal del *RB1* presentaron una LLA-T de tipo cortical ( $p= 0,002$ ), asociación descrita anteriormente en la LLA-T<sup>112,505</sup>. En cambio, solo 1/9 pacientes con deleción en la región del brazo 13q no focal del *RB1*, presentó una LLA-T cortical, sugiriendo que los dos tipos de deleciones tenían una repercusión diferente en los blastos. Por tanto, las deleciones del *RB1* focal y las *RB1* no focales se consideraron como grupos de alteraciones diferentes.

Hubo 14 pacientes con deleciones de tamaño variable del brazo 16q, pero todas incluyeron el gen *CTCF*. En este caso también se observaron dos patrones de alteraciones (Fig. 30): deleciones de más de 30 Mb que afectaron a la práctica totalidad del brazo 16q, y deleciones inferiores a 5 Mb que afectaron de un modo más focal a la región donde se localiza el gen *CTCF*. En este caso, no se observaron diferencias entre el inmunofenotipo de los pacientes con deleciones de mayor tamaño frente a los de menor tamaño, siendo el subtipo LLA-T madura el más común de los pacientes con estas alteraciones (OR deleciones *CTCF* - inmunofenotipo madura: 3,8;  $p=0,03$ ). Estos datos sugirieron que, independientemente del tamaño de la alteración, la presencia de deleciones que incluyeron el gen *CTCF* identificaban a pacientes con características similares.



**Figura 29. Delecciones del gen *RB1*.** Tamaño y posición de las delecciones que incluyen al gen *RB1*. Los distintos tipos de delecciones, agrupadas en función de si incluyen o no el *RB1*, se representan en distinto color. Para cada alteración se indica el identificador del paciente (TALL).

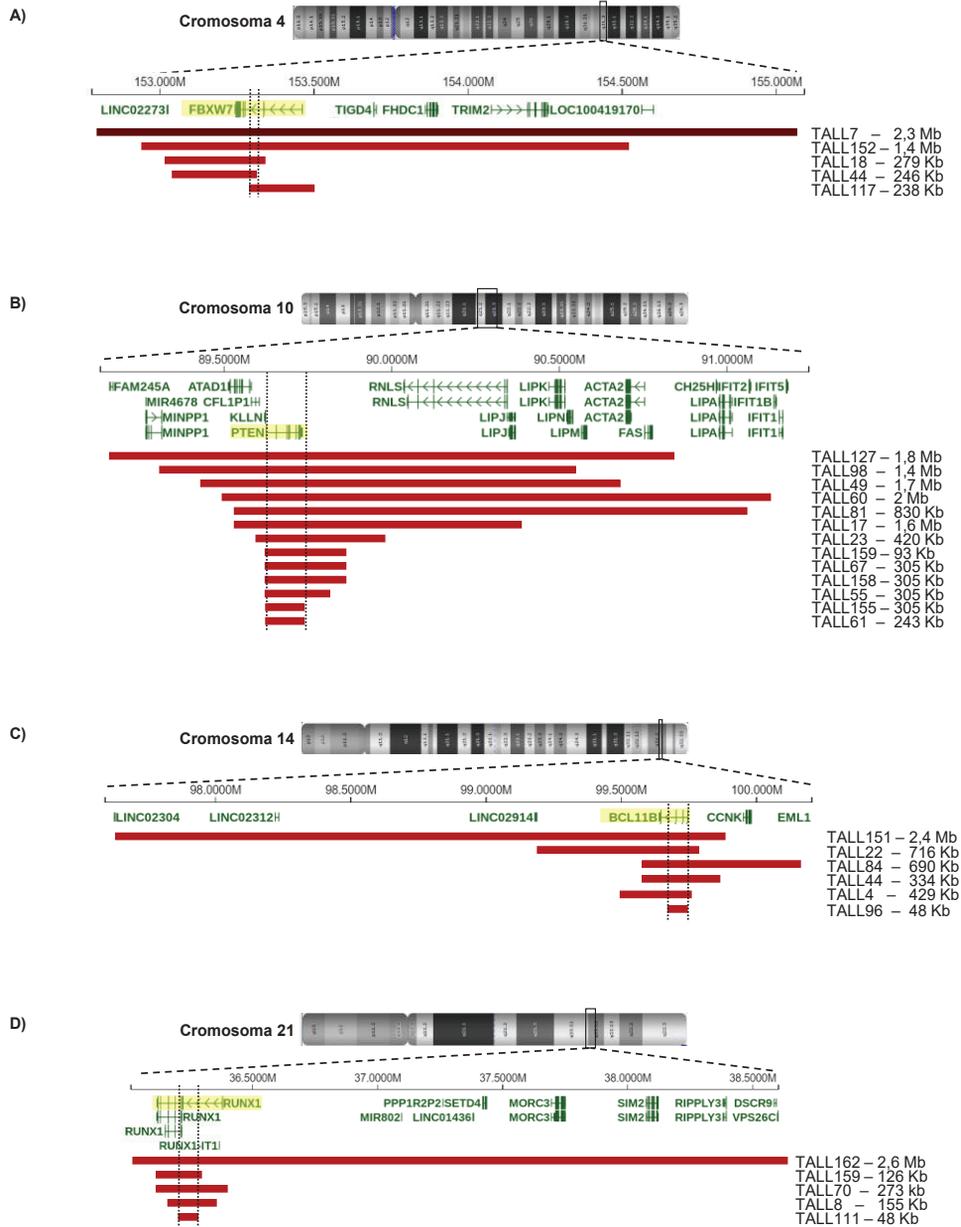


**Figura 30. Delecciones del gen *CTCF*.** Tamaño y posición de las diferentes delecciones que incluyen el gen *CTCF*. Para cada alteración se indica el identificador del paciente (TALL).

- **Alteraciones en las que la mínima región deleccionada identifica a un gen *driver***

Hubo otro conjunto de alteraciones con tamaños variables pero que no se asociaron con un inmunofenotipo específico y no se pudieron establecer grupos en base a esto. No obstante, presentaron una mínima región alterada común que coincidió con un gen *driver* de la LLA-T, sugiriendo que este podría ser el gen diana de la alteración. Fue el caso de las delecciones que afectaron al inicio del brazo 1p (Fig. 31A), donde la mínima región alterada coincidió con la localización del gen *RPL22*, y las que afectaron al final de este mismo brazo (Fig. 31B) que coincidió con la localización del gen *RPL5*. Las delecciones que afectaron a las citobandas 16p13.3-p13.11, la mínima región alterada incluyó el gen *CREBBP* (Fig. 31C). También fueron recurrentes las delecciones en el brazo 17p donde la mínima región alterada coincidió con el gen *TP53* (Fig. 31D). Finalmente, hubo un conjunto de delecciones que, aunque con un tamaño variable, todas fueron menores a 2,5 Mb y la mínima región alterada también coincidió con un gen *driver* de la LLA-T. Fue el caso de las delecciones localizadas en la citobanda 4q31.3 donde se encuentra *FBXW7* (Fig. 32A), las que afectaron a las citobandas 10q23.2-q23.31 donde se encuentra *PTEN* (Fig. 32B), las localizadas en la citobanda 14q32.2 donde se ubica *BCL11B* (Fig. 32C), y las que afectaron a la citobanda 21q22.12 donde se encuentra *RUNX1* (Fig. 32D).

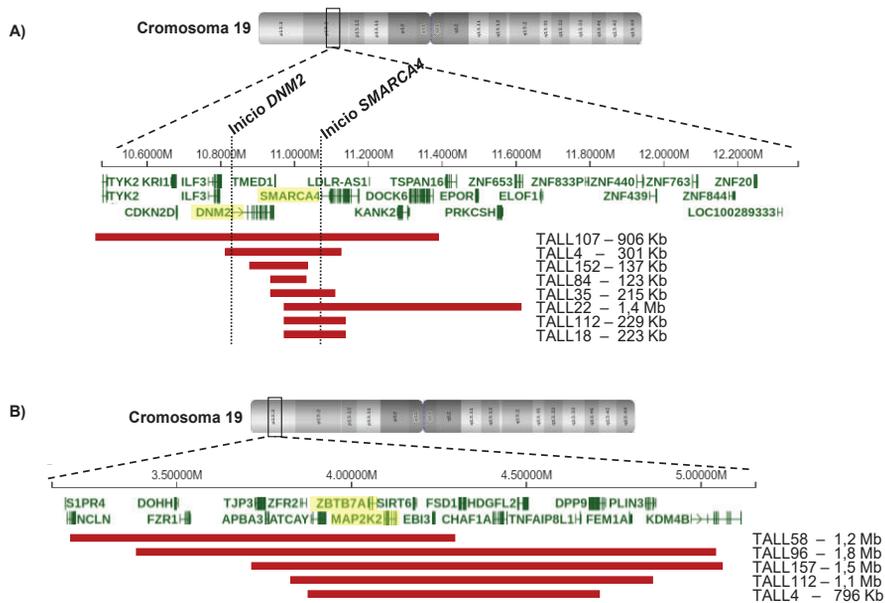




**Figura 32. Alteraciones < 2,5 Mb mínima región delecionada identifican a un gen driver.** A) Deleciones citobanda 4q31.3 que convergen en el gen *FBXW7*. B) Deleciones citobandas 10q23.2-q23.31 que convergen en el gen *PTEN*. C) Deleciones citobanda 14q32.2 que convergen en el gen *BCL11B*. D) Deleciones citobanda 21q22.12 que convergen en el gen *RUNX1*. La mínima región común se encuentra delimitada por dos líneas discontinuas y el gen *driver* resaltado en amarillo.

- **Alteraciones en las que la mínima región alterada no identifican a un gen *driver***

Hubo dos regiones dentro del cromosoma 19 con deleciones de tamaño variable y que incluyeron más de un gen *driver* de la LLA-T, no siendo posible determinar el gen diana, en el caso de que lo hubiese. Fue el caso de las deleciones en la citobanda 19p13.2, que incluyen dos genes *drivers* de la LLA-T (*DNM2* y *SMARCA4*) separados por menos de 150 kb (Fig. 33A). Esta alteración se estudió como deleción de la citobanda 19p13.2. También ocurrió lo mismo en las deleciones de la citobanda 19p13.3 donde la mínima región común delecionada incluyó los genes *ZBTB7A* y *MAP2K2* (Fig. 33B). Esta alteración se estudió como deleción de la citobanda 19p13.3.



**Figura 33. Alteraciones mínima región alterada no identifican a un gen *driver*.** A) Mínima región común de las deleciones localizadas en la citobanda 19p13.2. Se marca con líneas discontinuas el inicio de los genes *drivers* de la LLA-T incluidos en esta región. B) Mínima región común de las deleciones de la citobanda 19p13.3, se marca en amarillo los posibles genes *drivers* de esta región. Para cada alteración se indica el tamaño y el identificador del paciente (TALL).

- **Alteraciones focales de genes *drivers* de la LLA-T**

Hubo un conjunto de deleciones que afectaron únicamente a genes específicos involucrados en la patogénesis de la LLA-T. Fue el caso de las deleciones focales de los genes *TET2*, *LEF1*, *CDK6*, *WT1*, *ELF1*, *PTPN2*, *PHF6* y de la duplicación focal de *MYB* (Tabla 14).

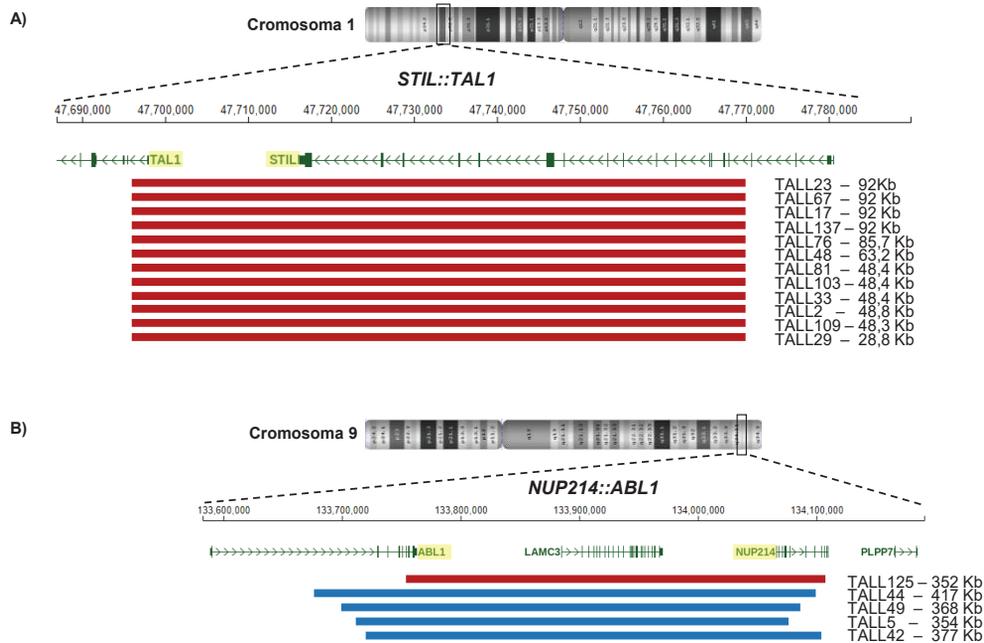
**Tabla 14. Alteraciones focales**

Alteración	ID	Tipo	Cromosoma	Citobanda	Coordenadas	Tamaño (bp)
<i>TET2</i>	TALL64	Delección	4	q24-q24	105780243-106328329	548.086
	TALL93	Delección		q24-q24	105867407-106380469	513.062
<i>LEF1</i>	TALL62	Delección	4	q25-q25	108979954-109084557	104.603
	TALL109	Delección		q25-q25	108991644-109098953	107.309
	TALL89	Delección		q25-q25	109019239-109084557	65.318
	TALL97	Delección		q25-q25	109019239-109084557	65.318
	TALL151	Delección		q25-q25	109019239-109084557	65.318
	TALL94	Delección		q25-q25	109022906-109084557	61.651
	TALL84	Delección		q25-q25	109041241-109084557	43.316
	TALL155	Delección		q25-q25	109062645-109084557	21.912
<i>CDK6</i>	TALL159	Delección	7	q21.2-q21.2	92234736-92480847	246.111
	TALL3	Delección		q21.2-q21.2	92234736-92470489	235.753
	TALL117	Delección		q21.2-q21.2	92235024-92480915	245.891
	TALL109	Delección		q21.2-q21.2	92243718-92493844	250.126
	TALL73	Delección		q21.2-q21.2	92243811-92480682	236.871
	TALL158	Delección		q21.2-q21.2	92243811-92480682	236.871
	TALL70	Delección		q21.2-q21.2	92266275-92480683	214.408
	TALL107	Delección		q21.2-q21.2	92366255-92480884	114.629
<i>WT1</i>	TALL49	Delección	11	p14.1-p13	24712000-25500679	788.679
	TALL118	Delección		p13-p13	32164176-32428040	263.864
	TALL34	Delección		p13-p13	32294646-32556376	261.730
	TALL5	Delección		p13-p13	32308361-32662850	354.489
<i>ELF1</i>	TALL70	Delección	13	q14.11-q14.11	41068436-41848458	780.022
	TALL159	Delección		q14.11-q14.11	41500549-41604107	103.558
	TALL73	Delección		q14.11-q14.11	41500550-41608111	107.561
	TALL158	Delección		q14.11-q14.11	41509092-41642414	133.322
	TALL61	Delección		q14.11-q14.11	41548370-41593536	45.166
<i>PTPN2</i>	TALL111	Delección	18	p11.21-p11.21	12779946-12906492	126.546
	TALL78	Delección		p11.21-p11.21	12779946-12906492	126.546
	TALL44	Delección		p11.21-p11.21	12779946-12903307	123.361
	TALL9	Delección		p11.21-p11.21	12779946-12894919	114.973
	TALL146	Delección		p11.21-p11.21	12779946-12894919	114.973
	TALL35	Delección		p11.21-p11.21	12779946-12882107	102.161
	TALL60	Delección		p11.21-p11.21	12779946-12882107	102.161
	TALL96	Delección		p11.21-p11.21	12781377-12882107	100.730
<i>PHF6</i>	TALL146	Delección	X	q26.2-q26.2	133544559-133587838	43.279
<i>MYB</i>	TALL28	Ganancia	6	q23.3-q23.3	135409899-135675217	265.318
	TALL73	Ganancia		q23.3-q23.3	135417402-135675162	257.760
	TALL32	Ganancia		q23.3-q23.3	135463996-135683603	219.607
	TALL40	Ganancia		q23.3-q23.3	135467724-135680880	213.156
	TALL82	Ganancia		q23.3-q23.3	135471715-135756666	284.951

	TALL13	Ganancia		q23.3-q23.3	135471715-135680602	208.887
	TALL78	Ganancia		q23.3-q23.3	135482565-135740496	257.931
	TALL47	Ganancia		q23.3-q23.3	135486125-135727064	240.939
	TALL38	Ganancia		q23.3-q23.3	135488016-135709118	221.102
	TALL62	Ganancia		q23.3-q23.3	135492128-135668993	176.865

• **Alteraciones que formaron genes de fusión**

Se detectaron alteraciones que sugirieron la formación de genes de fusión. Fue el caso de las deleciones que dieron lugar a la fusión *STIL::TAL 1* (Fig. 34A) y las alteraciones que dieron lugar a la fusión *NUP214::ABL 1* (Fig. 34B).



**Figura 34. Alteraciones que formaron genes de fusión.** A) Formación del gen de fusión *STIL::TAL 1*. B) Formación del gen de fusión *NUP214::ABL 1*. Las deleciones se indican con segmentos de color rojo y las duplicaciones con segmento de color azul. Los genes afectados se marcan en amarillo. Para cada alteración se indica el identificador del paciente (TALL).

**1.3. Panorama mutacional de la LLA-T**

Con el fin de identificar los genes implicados en la patogénesis de la LLA-T, se agruparon los genes afectados por alteraciones tipo SNV/indel con los genes diana de las alteraciones tipo CNV, identificados según se ha explicado en el apartado anterior.

Se consideraron aquellos genes alterados en al menos 5 pacientes, identificándose una mediana [extremos] de 5 [1-13] alteraciones/paciente para los casos con dato de TDS y SNP-*arrays* (n=133). Globalmente, fueron frecuentes las mutaciones en genes involucrados en vías de señalización implicadas en la regulación del desarrollo de las células T (Fig. 35).

En concreto, en la vía de señalización NOTCH y la vía de señalización JAK/STAT. En el primer caso, se identificaron alteraciones en los genes *NOTCH1* (62%) y *FBXW7* (21%). La vía de señalización NOTCH1 es la más frecuentemente mutada en la LLA-T y, en el caso de nuestra cohorte, el 68% de los pacientes presentaron alteraciones en estos genes. Respecto a la vía de JAK/STAT, se identificaron mutaciones en *JAK3* (13%), *JAK1* (8%), *DNM2* (8%), *IL7R* (7%) y *PTPN2* (6%). La vía de señalización RAS fue otra de las mutadas en la cohorte de estudio, donde se identificaron mutaciones en los genes *NRAS* y *KRAS* (11%) que se consideraron como único gen por presentar una estructura y función homóloga.

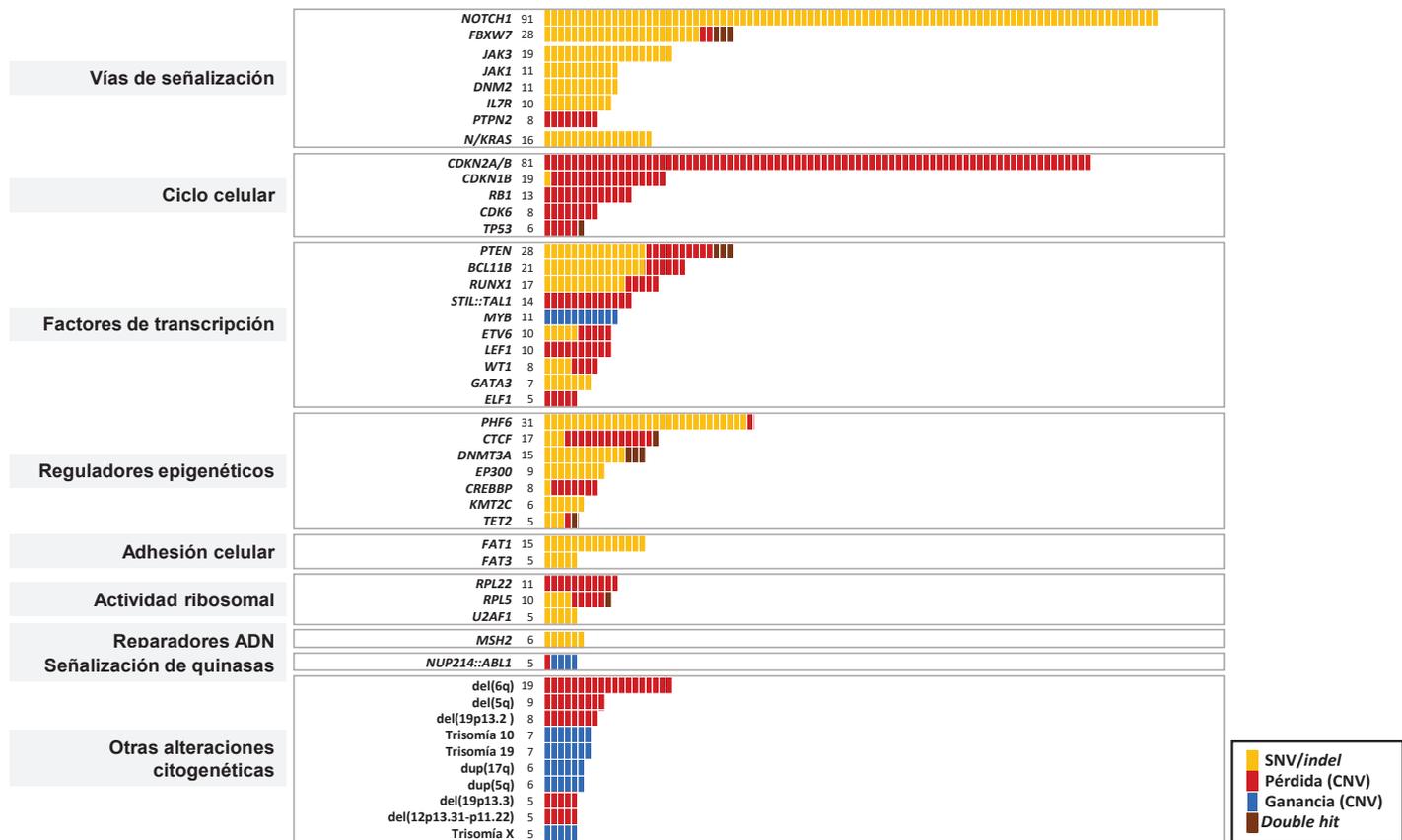
También se encontraron mutaciones en genes involucrados en la regulación del ciclo celular (*CDKN2A/B* (60%), *CDKN1B* (14%), *RB1* (10%), *CDK6* (6%), *TP53*(5%)), donde en este caso la mayoría de las alteraciones fueron deleciones.

Las alteraciones en genes que codifican a factores de transcripción que actúan como supresores tumorales (*PTEN*(20%), *BCL11B*(16%), *RUNX1*(13%), *ETV6*(7%), *LEF1*(8%), *WT1*(6%), *GATA3*(5%)) o como oncogenes (*MYB*(8%), *STIL::TAL1*(10%)) también fueron recurrentes; como también lo fueron las que afectaron a genes que actúan como reguladores epigenéticos (*PHF6*(23%), *CTCF*(13%), *DNMT3A*(10%), *EP300*(6%), *CREBBP*(6%), *KMT2C*(4%), *TET2*(4%)); genes involucrados en la adhesión celular (*FAT1*(10%) y *FAT3*(4%)); genes reguladores de la actividad ribosomal (*RPL22*(8%), *RPL5*(7%) y *U2AF1*(3%)); y genes con funciones reparadoras del ADN (*MSH2*(4%)).

Además, hubo pacientes que presentaron alteraciones citogenéticas que dieron lugar al gen de fusión *NUP214::ABL1* (4%) descrito por ser una forma de activación de la tirosina quinasa *ABL1*<sup>320</sup>.

Finalmente, también hubo alteraciones citogenéticas frecuentes. Fue el caso de las del(6q) (14%) y del(5q) (7%); la deleción de las citobanda 19p13.2 (6%) y de la citobanda 19p13.3 (4%); las trisomías del cromosoma 10 (5%), 19 (5%) y X (4%); las dup(17q) (5%) y dup(5q) (5%); y la deleción de la región del brazo 12p que comprendió las citobandas p13.31-p11.22 (4%) y que fue la región que incluyó los genes *ETV6*, *CDKN1B* y *KRAS*. Como se observa en la figura 35, hubo genes exclusivamente afectados por SNV/indels, otros solo afectados por deleciones y finalmente otros afectados por ambos tipos de alteraciones.

Hubo pacientes que presentaron combinaciones de dos tipos de alteraciones en un único gen, a modo de *double hit*. Fue el caso de los genes *DNMT3A*, *PTEN* y *FBXW7*, que presentaron SNV/indel y Del/LOH en 3 pacientes; y los genes *ETV6*, *CDKN1B*, *TET2* y *TP53* que cada uno se encontró mutado y delecionado en un paciente, sugiriendo que estos genes requieren de dos eventos para desarrollar la LLA-T.



**Figura 35. Panorama mutacional de la cohorte de estudio descriptiva (n=145).** Se muestran las alteraciones que fueron recurrentes en al menos 5 pacientes. Las alteraciones se agrupan según su implicación en vías de señalización, función sobre la actividad celular y finalmente las alteraciones citogenéticas. *Double hit* genes que presentaron en el mismo paciente dos tipos de alteraciones (SNV/deleción o SNV/LOH).

#### 1.4. Cooperación entre las alteraciones genéticas en el desarrollo de la LLA-T y en la definición de las características biológicas de los pacientes

Dada la gran heterogeneidad genética de la LLA-T, se buscaron asociaciones entre parejas de genes que se encontraron mutados en los mismos pacientes con el fin de evaluar la interacción entre los diferentes genes o la cooperación entre diferentes vías de señalización (Fig. 36A).

De este modo, se observó que los genes *NOTCH1*, *RB1*, *BCL11B* y *CDKN2A/B* estuvieron alterados en los mismos pacientes en diferentes combinaciones. Las alteraciones en *BCL11B* co-ocurrieron junto a las mutaciones en *NOTCH1* (OR= 169,4; q= 0,003), las deleciones en *CDKN2A/B* (OR= 16,7; q= 0,01) y las deleciones en *RB1* (OR= 13,2; q= 0,006). Por su parte, *RB1* mostró una tendencia a estar alterado en los mismos pacientes que presentaron mutaciones en *NOTCH1* (OR= 92,9; q= 0,06) y los pacientes que presentaron deleciones en *CDKN2A/B* (OR= 99,4; q= 0,06).

Las deleciones de *CDKN2A/B* fueron mutuamente excluyentes con las mutaciones en *DNMT3A* (OR=0,09; q= 0,03) y mostraron una tendencia a serlo con las alteraciones en *RUNX1* (OR= 0,17; q= 0,06).

Los pacientes que presentaron el gen de fusión *STIL::TAL1*, mostraron una tendencia a tener deleciones en el brazo 6q (OR= 7,1; q= 0,07). Hubo una tendencia positiva entre las mutaciones en *FBXW7* y *FAT1* (OR= 4,56; q=0,17).

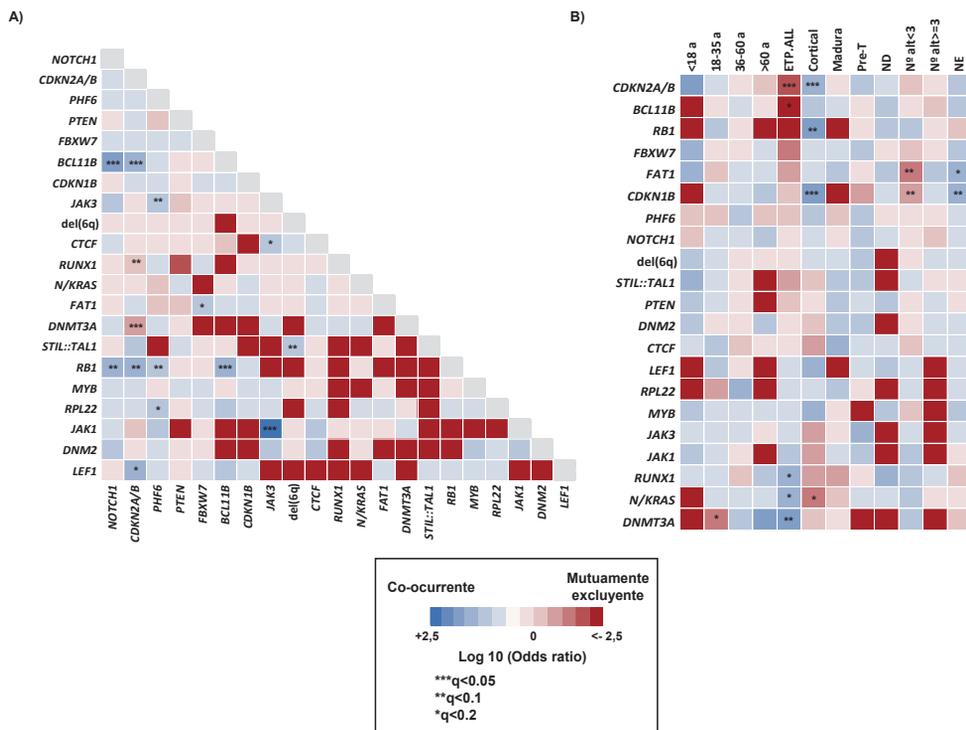
Finalmente, los genes *JAK3* y *JAK1* se encontraron mutados en los mismos pacientes (OR= 114, q= 0,0002), a su vez el gen *JAK3* mostró una tendencia a estar mutado en los pacientes que presentaron mutaciones en *PHF6* (OR= 5; q= 0,07) y en los pacientes con mutaciones en *PHF6* también se encontraron alteraciones en *RPL22* (OR=5,46; q=0,17), siendo esta relación una tendencia.

Por otro lado, también se buscaron asociaciones entre las características biológicas de los pacientes (edad, inmunofenotipo y número de alteraciones en el cariotipo) y los genes alterados presentes en los mismos (Fig. 36B).

De este modo, los pacientes del subtipo cortical presentaron mutaciones en *CDKN1B* (OR= 8,4; q= 0,02), deleciones en *CDKN2A/B* (OR= 4,4; q= 0,03) y una tendencia a tener deleciones en *RB1* (OR=7,2; q= 0,08). En cambio, las deleciones en *CDKN2A/B* fueron mutuamente excluyentes con el subtipo ETP-ALL (OR= 0,04, q=6,8 X 10<sup>-4</sup>), al igual que las mutaciones en *BCL11B* (OR=0; q= 0,17), aunque esta última relación fue una tendencia.

El subtipo ETP-ALL se caracterizó por presentar mutaciones en *DNMT3A* (OR= 7,5; q= 0,08) y *N/KRAS* (OR= 4,6; q= 0,15), siendo ambas asociaciones una tendencia.

Finalmente se observó que los pacientes con menos de 3 alteraciones en el cariotipo tuvieron tendencia a no presentar mutaciones en *FAT1* (OR=0,15; q= 0,07) ni alteraciones en *CDKN1B* (OR= 0,16; q= 0,08). En cambio, las mutaciones en este último gen fueron frecuentes en los pacientes con un cariotipo no evaluable (OR= 5,2; q= 0,07), siendo esta asociación una tendencia.



**Figura 36. Asociaciones entre pares de genes y entre genes y características biológicas.** A) Pares de asociaciones entre genes co-mutados o mutuamente excluyentes en el mismo paciente. B) Pares de asociaciones entre genes co-ocurrentes o mutuamente excluyentes con las características biológicas del paciente. Las asociaciones positivas están marcadas de tonalidad azul mientras que las negativas de granate. a: años; ND: inmunofenotipo no determinado; alt: alteraciones; NE: cariotipo no evaluable.

### 1.5. Evaluación de la jerarquía clonal de la LLA-T

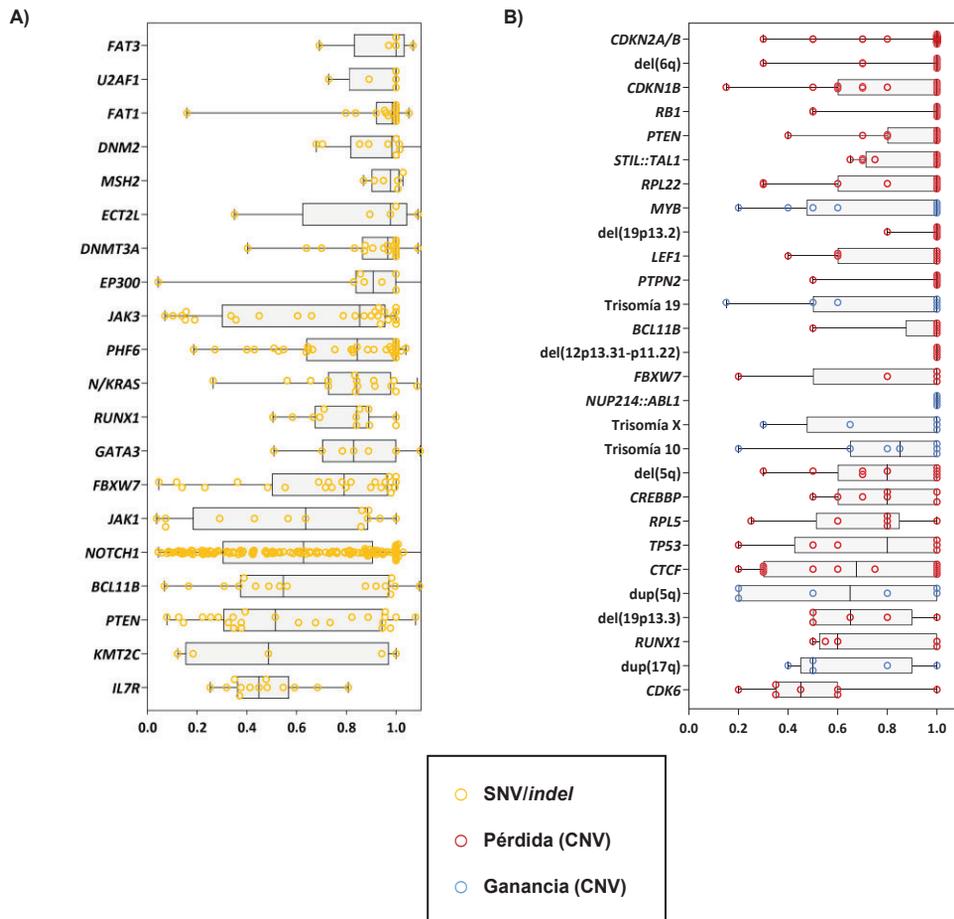
Una aproximación para evaluar el orden de adquisición de las alteraciones genéticas en el desarrollo de la LLA-T es mediante el estudio de la VAF/CCF de cada alteración que aporta una estimación del porcentaje de células afectadas por cada alteración.

Con este fin, las VAF de todas las alteraciones se corrigieron por el porcentaje de blastos y la presencia de CNV, según se ha explicado en la metodología. Se consideraron subclonales aquellas alteraciones con una CCF <50% (VAF < 25%).

Todos los pacientes presentaron al menos una alteración clonal (SNV/indel o CNV) acompañada de una o varias alteraciones subclonales. Las alteraciones tipo SNV/indel con mayor CCF (mediana de la CCF >80%) fueron: *FAT1*, *DNMT2*, *MSH2*, *ETC2L*, *DNMT3A*, *EP300*, *JAK3*, *PHF6*, *N/KRAS*, *RUNX1*, *GATA3*; lo que indicaría que se tratan de eventos tempranos. Mientras que la mayoría de los genes con una menor CCF (mediana CCF < 80%) fueron genes involucrados en las vías de señalización más frecuentemente alteradas en la LLA-T (*FBXW7*, *JAK1*, *NOTCH1*, *BCL11B*, *PTEN*, *KMT2C*, *IL7R*), y estaría indicando que se tratan de eventos más tardíos en el desarrollo de la misma (Fig. 37A).

En el caso específico de *NOTCH1* se observó gran variabilidad: la variante más subclonal presentó una CCF de 4% y la más clonal del 100%, siendo la mediana 63,3%.

Para las alteraciones tipo CNV también se estimó la CCF. En este caso, la mayoría de las alteraciones presentaron una mediana de la CCF >80%, con la excepción de las deleciones en *CTCF*, la dup(5q), la deleción de la citobanda 19p13.3, las deleciones en *RUNX1*, la dup(17q) y las duplicaciones en *CDK6* (Fig. 37B). Se observó que la mediana de la CCF de todas las alteraciones tipo CNV fue mayor que la de las alteraciones SNV/indel (1 vs 0,83;  $p < 0,0001$ ).



**Figura 37. CCF de las alteraciones tipo SNV/indel y CNV.** A) Gráfico tipo *box and whisker* (caja y bigotes) de la CCF de cada SNV/indel identificada para cada gen. Cada alteración se identifica con un círculo amarillo. B) Gráfico tipo *box and whisker* (caja y bigotes) de la CCF de cada CNV, cada pérdida se identifica con un círculo rojo y cada ganancia con un círculo azul.

También hubo diferencias entre las medianas de las CCF de las pérdidas y las ganancias, siendo más alta la de las pérdidas (1 vs 0,8;  $p=0,0002$ ). Los genes *PTEN*, *FBXW7*, *BCL11B* y *RUNX1* se presentaron alterados de manera recurrente por SNV/indel y por CNV. Las CCF de *PTEN* y *BCL11B* fueron significativamente mayor para el caso del CNV que para el caso de SNV ( $p=0,001$ ,  $p=0,0130$ ; respectivamente); mientras que no hubo diferencias para el caso de *FBXW7* y *RUNX1*.

Del total de genes analizados, *TP53* y *N/KRAS* son especialmente frecuentes en la recaída, y en algunos trabajos se ha mostrado la presencia de mutaciones subclonales en estos genes en el diagnóstico que se expanden en la recaída<sup>414,416,417</sup>. No obstante, en este estudio la mediana de la CCF de las alteraciones para ambos genes fue superior al 50% y solo un paciente presentó una alteración subclonal para el caso de *TP53* (CCF=20%), y otro paciente una variante subclonal para el caso de *N/KRAS* (CCF= 26%).

## RESULTADOS DEL OBJETIVO 2

### 1. Impacto pronóstico de las alteraciones detectadas por TDS y SNP-arrays

Del total de los 146 pacientes de LLA-T analizados por TDS y 134 por SNP-arrays se seleccionaron aquellos pacientes adultos de los cuales se tenía la información clínica, biológica y de pronóstico completa y que habían sido tratados de manera homogénea según los protocolos PETHEMA para LLA: AR-11 y AR-03, ambos con seguimiento de la enfermedad residual por citometría de flujo. En total supusieron 116 pacientes analizados por TDS y 107 analizados por SNP-arrays. Los marcadores pronósticos a analizar fueron la probabilidad (a partir ahora prob.) de OS y la CIR asociadas a las alteraciones recurrentes en al menos 5 pacientes.

La mediana de edad de la cohorte de estudio analizada por TDS (n=116) fue de 37 [16-61] años, siendo una mayoría de pacientes hombres (74% vs 26%). En cuanto a los datos clínicos y de respuesta al tratamiento (Tabla 15), un 83% de los pacientes tuvieron una ERM < 0,1% tras el primer bloque de inducción y un 88% alcanzaron la remisión completa (RC) tras el primer o segundo bloque de inducción. La mayoría de los pacientes, un 72%, siguieron un tratamiento de quimioterapia tras los bloques de consolidación frente a un 28% que fueron trasplantados. La OS a los 5 años fue del 36% y la CIR a los 5 años fue del 54%.

**Tabla 15. Características biológicas, respuesta al tratamiento y pronóstico cohorte de estudio analizada por TDS (n=116)**

Características biológicas	
Mediana de edad al diagnóstico [extremos], años	37 [16-61]
Género	
Hombres, n (%)	86 (74%)
Mujeres, n (%)	30 (26%)
Características clínicas	
ERM <0,1% a día +35	74/ 89 (83%)
RC tras inducción-1 o inducción-2	102/116 (88%)
Tratamiento post-consolidación	
Quimioterapia	59/82 (72%)
Alo-TPH	23/82 (28%)
Probabilidad de Supervivencia Global 5 años [IC95%]	36% [26%-46%]
Incidencia acumulada de recaída 5 años [IC95%]	54% [43%-64%]

ERM: Enfermedad residual medible; RC: Remisión completa; Alo-TPH: Trasplante de progenitores hematopoyéticos alogénico.

#### 1.1. Impacto en la supervivencia global de las alteraciones tipo SNV/indel

En primer lugar, se evaluó el impacto de las alteraciones tipo SNV/indel detectadas por TDS en al menos 5 pacientes en la OS de los pacientes con datos clínicos completos (n=116).

### 1.1.1. Impacto en la supervivencia global de las mutaciones que afectan a genes de la vía de señalización de NOTCH1 (*NOTCH1* y *FBXW7*)

*NOTCH1* fue el gen más mutado en esta cohorte (60% de los pacientes presentaron mutaciones en *NOTCH1*) y en la LLA-T en general, y por esta razón el análisis de su impacto pronóstico tiene especial interés. Además, las mutaciones en *NOTCH1* se han asociado con una OS más prolongada en estudios anteriores<sup>237,251</sup>, aunque otros estudios no han validado estos resultados<sup>236,248</sup>.

En este trabajo en concreto, no se confirmó dicha asociación pues los pacientes no mutados en *NOTCH1* presentaron una prob. OS a los 5 años de 35% [IC 95%: 23%-47%] frente a los mutados que fue del 36% [IC95%: 16%-56%] ( $p=0,313$ ) (Fig 38A).

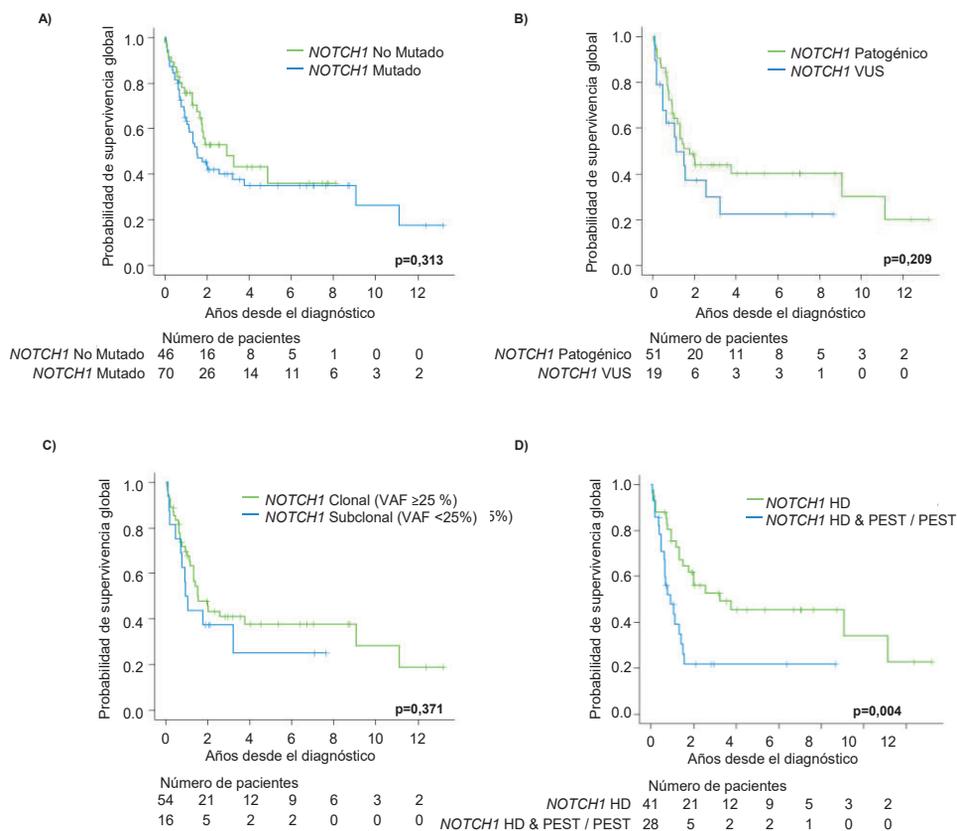
Se analizaron en más detalle estas mutaciones, comparando la prob. OS de los pacientes que presentaron únicamente mutaciones en *NOTCH1* categorizadas como VUS (prob. OS a los 5 años 22% [IC95: 1%-43%]) frente aquellos que presentaron mutaciones categorizadas como patogénicas (prob. OS a los 5 años 40% [IC: 25% -55%]), con el fin de comprobar si ambos tipos de categorizaciones podrían tener un efecto diferente sobre la supervivencia. No obstante, tampoco se obtuvieron diferencias significativas ( $p=0,243$ ). (Fig 38B).

También se evaluó las diferencias de OS entre los pacientes que únicamente presentaron mutaciones en *NOTCH1* subclonales, consideradas como tales aquellas cuya VAF fue menor al 25%, (prob. OS a los 5 años 25% [IC95%: 0%-50%]) frente a aquellos que tuvieron mutaciones clonales ( $VAF \geq 25\%$ ) (prob. OS a los 5 años 38% [IC95%: 24%-52%]), sin haber diferencias significativas ( $p=0,380$ ) (Fig 38C).

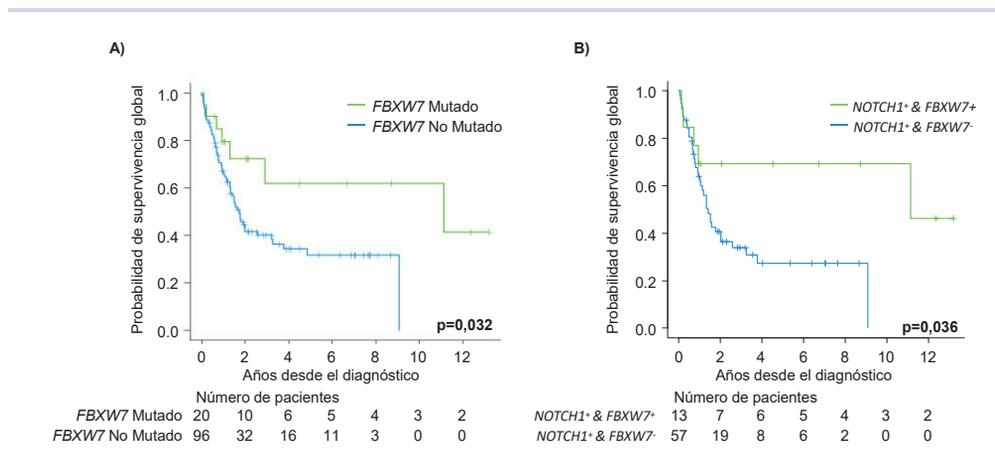
Por último, se analizó si el dominio afectado de *NOTCH1* (HD o PEST) podía estar correlacionado con la supervivencia, pues el mecanismo por el cual NOTCH1 permanece activo cuando el gen está mutado es diferente en función del dominio afectado<sup>506</sup>. Del total de pacientes con mutaciones en *NOTCH1*, 69/70 presentaron mutaciones en los dominios HD o PEST, mientras que en el restante la mutación afectó al dominio EGF-like. De este modo, 41/69 (60%) pacientes presentaron mutaciones solo en el dominio HD, 19/69 (28%) en ambos dominios (HD + PEST) y 9/69 (13%) solo en el dominio PEST. Estudiando la OS de estos tres grupos de pacientes se observó que los pacientes con mutaciones en el dominio PEST y con mutaciones en el dominio HD + PEST tuvieron un pronóstico similar. Por esa razón estos dos grupos se evaluaron conjuntamente y se compararon con los pacientes que solo presentaron mutaciones en el dominio HD, observándose que el pronóstico de los pacientes con mutaciones en PEST y PEST + HD fue peor que el de los pacientes con mutaciones solo en el dominio HD (Fig. 38D).

A diferencia del gen *NOTCH1*, la presencia de mutaciones en *FBXW7* tuvo un impacto en la OS de los pacientes. En concreto, los pacientes con mutaciones en *FBXW7* presentaron una OS más prolongada que los pacientes sin mutaciones (prob. OS a los 5 años 62% [IC95: 36%-88%] frente a 32% [IC95: 21%-43%],  $p= 0,032$ ) (Fig. 39A), siendo además el único gen mutado asociado a mejor supervivencia. En cambio, este mejor pronóstico no se mantuvo al considerar los pacientes con mutaciones en *NOTCH1* y/o *FBXW7* ( $n=77$ ).

Sin embargo, los pacientes con *NOTCH1* y *FBXW7* mutado (n=13) presentaron mayor probabilidad de supervivencia (prob. OS a los 5 años 69% [IC95%: 37%-87%]) que los pacientes con *NOTCH1* mutado y sin mutaciones en *FBXW7* (n=57) (prob. OS a los 5 años 27% [IC95%: 15% -41%]) (p=0,036), mostrando el efecto positivo de las mutaciones de *FBXW7* sobre la OS de los pacientes. (Fig. 39B).



**Figura 38. Impacto en la supervivencia global a los 5 años de las mutaciones en *NOTCH1*.** A) OS en función de la presencia de mutaciones en *NOTCH1*. La OS estimada para los pacientes con mutaciones en *NOTCH1* fue 35% [IC95%: 23%-47%] y 36% [IC95%: 16%-56%] para los no mutados. B) OS en función del impacto funcional de las variantes de *NOTCH1*. La OS estimada para los pacientes con mutaciones patogénicas fue 40% [IC95%: 25%-55%] y 22% [IC95%: 1%-43%], para los pacientes con mutaciones de significado incierto (*Variants of unknown significance*, VUS). C) OS según la clonalidad de las variantes. La OS estimada para los pacientes con mutaciones clonales (VAF ≥25%) fue 38% [IC95%, 24%-52%] y 25% [IC95%, 0%-50%] para los pacientes con mutaciones subclonales (VAF < 25%). D) OS según los dominios mutados de *NOTCH1*. La OS estimada para los pacientes con mutaciones en el dominio HD fue 45% [IC95%: 28%-61%] y 22% [IC95% 8%-40%] para los pacientes con mutaciones en el dominio PEST o PEST+HD.



**Figura 39. Impacto en la supervivencia global a los 5 años de las mutaciones en *FBXW7*.** A) OS en función de la presencia de mutaciones en *FBXW7*. La OS estimada para los pacientes con mutaciones en *FBXW7* fue 62% [IC95: 36%-88%] y 36% [IC95%: 16%-56%] para los no mutados. B) OS en función de la relación de mutaciones en *NOTCH1* ± *FBXW7*. La OS estimada para los pacientes con mutaciones en *NOTCH1* y *FBXW7* fue 69% [IC95%: 37%-87%] y 27% [IC95%: 15% -41%] para los pacientes con mutaciones en *NOTCH1* y sin mutaciones en *FBXW7*.

### 1.1.2. Las mutaciones en *DNMT3A*, *U2AF1*, *N/KRAS* y/o *MSH2* identifican a pacientes con peor supervivencia global

Del resto de genes analizados, se observó que los pacientes con mutaciones en *DNMT3A*, *N/KRAS*, *MSH2* y/o *U2AF1* presentaron una prob. OS menor que los pacientes sin estas mutaciones (Tabla 16).

**Tabla 16. Genes asociados a una peor supervivencia global**

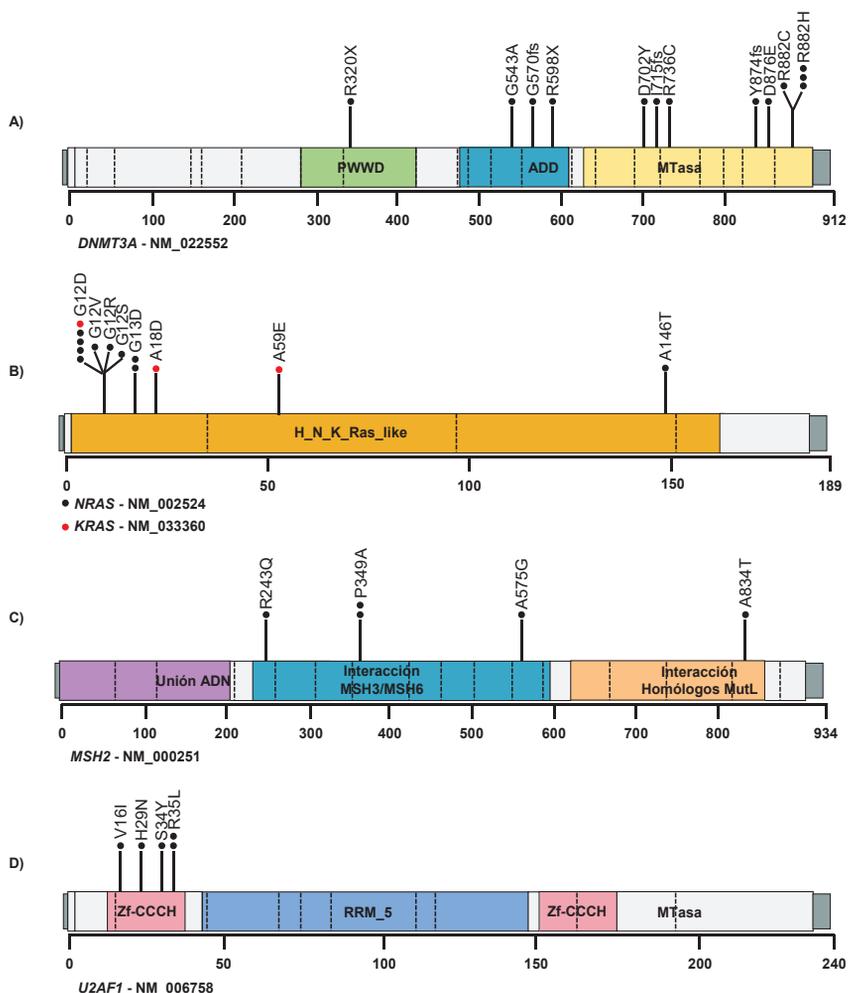
Gen	Número de pacientes (n=116)		Probabilidad de supervivencia global a los 5 años [IC 95%]		p
	Paciente con mutaciones	Paciente sin mutaciones	Paciente con mutaciones	Paciente sin mutaciones	
<i>DNMT3A</i>	10	106	13% [0%-37%]	38% [27%-49%]	0,001
<i>N/KRAS</i>	13	103	21% [0%-45%]	39% [28%-50%]	0,023
<i>MSH2</i>	5	111	20% [0%-55%]	37% [26%-48%]	0,036
<i>U2AF1</i> *	5	111	20% [0%-55%]	50% [40%-60%]	0,003

\*La probabilidad de supervivencia global fue estimada a los 2 años.

El gen *DNMT3A* estuvo mutado en 10 pacientes y tres de estos presentaron dos mutaciones. Todas las mutaciones afectaron a dominios funcionales, siendo el dominio metil-transferasa (MTasa) el más frecuentemente mutado, pues 9/13 mutaciones se situaron en esta región (Fig. 40A). Las mutaciones en *N/KRAS* afectaron a 13 pacientes y ninguno de ellos presentó más de una mutación en estos genes. Todas se situaron en el único dominio funcional de este gen (*H\_N\_K\_Ras\_like*), y 8/13 concretamente afectaron al aminoácido G12 (Fig. 40B). *MSH2* se presentó mutado en 5 pacientes, dos de ellos presentaron la misma mutación

(P349A). Todas las mutaciones se situaron en los dominios funcionales del gen, específicamente 4/5 afectaron al dominio de interacción con MSH3/MSH6 mientras que otra afectó al dominio de interacción con los genes homólogos de MutL (Fig 40C). Finalmente, las mutaciones en *U2AF1* afectaron a 5 pacientes. Dos pacientes presentaron la misma mutación (R35L) y todas ellas afectaron al mismo dominio (Zf-CCCH, dominio tipo *Zinc finger*) (Fig. 40D).

Teniendo una visión global de la respuesta al tratamiento y la evolución de estos pacientes, como se muestra en la figura 41, se observó que los pacientes con mutaciones en *DNMT3A*, *N/KRAS*, *MSH2* y/o *U2AF1* presentaron respuestas al tratamiento más lentas, realizaron un mayor número de ciclos de inducción, y murieron de forma prematura.



**Figura 40. Dominios proteicos afectados en los genes que identifican pacientes con una peor OS.** A) Localización de las mutaciones en *DNMT3A*. B) Localización de las mutaciones en *N/KRAS*. C) Localización de las mutaciones en *MSH2*. D) Localización de las mutaciones en *U2AF1*. En la parte inferior de cada gen se indica el transcrito de referencia (NM). Con un círculo se identifica cada uno de los pacientes que presenta la mutación descrita.