



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE MEDICINA

**ANÁLISIS TRANSCRIPTÓMICO PARA LA IDENTIFICACIÓN DE ALTERACIONES
GENÉTICAS ASOCIADAS A RECAÍDA Y MUERTE TEMPRANA DE PACIENTES
PEDIÁTRICOS CON LEUCEMIA LINFOBLÁSTICA AGUDA**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

M. EN C. BÁRCENAS LÓPEZ DIEGO ALBERTO

TUTORA PRINCIPAL DE TESIS: DRA. SILVIA JIMÉNEZ MORALES
INSTITUTO NACIONAL DE MEDICINA GENÓMICA, INMEGEN

COMITÉ TUTOR: DR. ENRIQUE RAMÓN ÁNGELES ANGUIANO
FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLÁN
DR. ALFREDO HIDALGO MIRANDA
INSTITUTO NACIONAL DE MEDICINA GENÓMICA, INMEGEN

CIUDAD UNIVERSITARIA, CDMX, 2023



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE MEDICINA

**ANÁLISIS TRANSCRIPTÓMICO PARA LA IDENTIFICACIÓN DE ALTERACIONES
GENÉTICAS ASOCIADAS A RECAÍDA Y MUERTE TEMPRANA DE PACIENTES
PEDIÁTRICOS CON LEUCEMIA LINFOBLÁSTICA AGUDA**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

M. EN C. BÁRCENAS LÓPEZ DIEGO ALBERTO

TUTORA PRINCIPAL DE TESIS: DRA. SILVIA JIMÉNEZ MORALES
INSTITUTO NACIONAL DE MEDICINA GENÓMICA, INMEGEN

COMITÉ TUTOR: DR. ENRIQUE RAMÓN ÁNGELES ANGUIANO
FACULTAD DE ESTUDIOS SUPERIORES CUAUTILÁN
DR. ALFREDO HIDALGO MIRANDA
INSTITUTO NACIONAL DE MEDICINA GENÓMICA, INMEGEN

CIUDAD UNIVERSITARIA, CDMX, 2023

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

OFICIO CPCB/039/2023

ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
P r e s e n t e

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **17 de octubre de 2022** se aprobó el siguiente jurado para el examen de grado de **DOCTOR EN CIENCIAS** del estudiante **BÁRCENAS LÓPEZ DIEGO ALBERTO** con número de cuenta **304116285** con la tesis titulada **“ANÁLISIS TRANSCRIPTÓMICO PARA LA IDENTIFICACIÓN DE ALTERACIONES GENÉTICAS ASOCIADAS A RECAÍDA Y MUERTE TEMPRANA DE PACIENTES PEDIÁTRICOS CON LEUCEMIA LINFOBLÁSTICA AGUDA”**, realizada bajo la dirección de la **DRA. SILVIA JIMÉNEZ MORALES**, quedando integrado de la siguiente manera:

Presidente: DRA. IRMA OLARTE CARRILLO
Vocal: DRA. MARÍA ANTONIETA CHÁVEZ GONZÁLEZ
Vocal: DR. ALEXANDER PEDROZA GONZÁLEZ
Vocal: DRA. ELISA MARÍA DORANTES ACOSTA
Secretario: DR. ENRIQUE RAMÓN ÁNGELES ANGUIANO

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
“POR MI RAZA HABLARÁ EL ESPÍRITU”
Ciudad Universitaria, Cd. Mx., a 10 de enero de 2023

COORDINADOR DEL PROGRAMA



DR. ADOLFO GERARDO NAVARRO SIGÜENZA



Agradecimientos institucionales:

- a) Al Posgrado en Ciencias Biológicas, UNAM
- b) Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada (CVU: 737534, No. de apoyo 717862).

También a CONACYT (IFC-2016-01-2119, PDCPN2013-01-215726, SALUD 2015-1-262190, CB-2015-1-258042, FONCICYT/37/2018), al Instituto Mexicano del Seguro Social (IMSS) (FIS/IMSS/PROT/1364, FIS/IMSS/PROT/1548, FIS/IMSS/PROT/895, FIS/IMSS/PROT/1533, FIS/IMSS/PROT/1782, FIS/IMSS/PROT/PRIO/14/031, FIS/IMSS/PROT/PRIO/15/048, FIS/IMSS/PROT/MD15/1504, FIS/IMSS/PROT/G15/1477), y al Instituto Nacional de Medicina Genómica (INMEGEN) (19/2019/I) por el financiamiento para realizar este proyecto.

- c) A mi tutora, Dra. Silvia Jiménez Morales, y los miembros de mi comité tutor, Dr. Enrique Ángeles Anguiano y Dr. Alfredo Hidalgo Miranda por sus aportaciones y mentoría para realizar este trabajo.

Dedicatoria

A mis padres, Lino y Rosaura

ÍNDICE

ÍNDICE DE TABLAS	
ÍNDICE DE FIGURAS	
ABREVIATURAS	
RESUMEN	1
ABSTRACT	3
ANTECEDENTES	5
Epidemiología del cáncer infantil	5
Etiología y origen de la LLA.....	7
Desarrollo prenatal de la LLA.....	9
Nuevo panorama genético de la LLA	12
Otros B	13
Alteraciones del gen <i>DUX4</i>	14
Rearreglos del gen <i>ZNF384</i> (<i>ZNF384-r</i>).....	16
Rearreglos del gen <i>MEF2D</i> (<i>MEF2D-r</i>).....	17
Fenocopias de subtipos recurrentes, Ph-like y <i>ETV6::RUNX1</i> -like.	18
Alteraciones conductoras de <i>PAX5</i> : <i>PAX5alt</i> y <i>PAX5-P80R</i>	20
Otras alteraciones del locus <i>IGH</i>	22
Nuevos subtipos de LLA-T.....	22
Asociación entre subtipos genéticos con edad, pronóstico y etnicidad	23
Situación epidemiológica de la LLA pediátrica en México y el mundo	26
PLANTEAMIENTO DEL PROBLEMA	30
PREGUNTA DE INVESTIGACIÓN	31
HIPÓTESIS.....	31
OBJETIVO GENERAL.....	32
Objetivos específicos.....	32
METODOLOGÍA	33
Selección de pacientes	33
Procesamiento de muestras para análisis de transcriptoma (RNA-seq).....	34
Preparación de las bibliotecas de secuenciación.....	35
Búsqueda de fusiones génicas con FusionCatcher	35
Búsqueda dirigida para identificación de fusiones del locus <i>IGH</i>	38
Identificación de perfil de expresión Ph-like con datos de microarreglos	39
Validación de fusiones génicas	40

Análisis estadístico	40
RESULTADOS	41
Características clínicas de la población analizada con RNA-seq	41
Pronóstico basado en características clínicas y clasificación de riesgo.....	43
Fusiones génicas identificadas en LLA	45
Identificación del perfil Ph-like con datos de microarreglos.....	48
Identificación de outliers con expresión de <i>CRLF2</i> y <i>DUX4</i>	49
Clasificación por subtipos genéticos en LLA-B.....	51
Características de las fusiones identificadas.....	54
MEF2D.....	54
PAX5-r	55
CRLF2-r	56
DUX4	59
ZNF384-r.....	63
Otras fusiones sin subtipo asignado en LLA.....	65
Características clínicas y pronóstico de los subtipos moleculares	67
DISCUSIÓN	73
CONCLUSIONES	83
REFERENCIAS BIBLIOGRÁFICAS	84
ANEXOS	103
Artículo de requisito para la obtención del grado	103
Otros artículos de investigación	121
Artículos de revisión.....	122
Tablas suplementarias.....	124
Congresos internacionales	132

ÍNDICE DE TABLAS

Tabla 1. Características clínicas y moleculares de población de estudio.....	42
Tabla 2. Variables clínicas asociadas a subtipos genéticos en LLA.....	70
Tabla 3. Sobrevida de subtipos genéticos en LLA.....	72

ÍNDICE DE FIGURAS

Figura 1. Desarrollo de la LLA de linaje B.....	9
Figura 2. Subtipos genéticos en LLA.....	13
Figura 3. Cinasas activadas en LLA Ph-like.....	20
Figura 4. Distribución de subtipos genéticos en LLA-B respecto al riesgo y edad.....	25
Figura 5. Tasas crudas de incidencia de LLA por edad respecto a la etnicidad.....	29
Figura 6. Fases del tratamiento de LLA.....	35
Figura 7. Conceptos básicos de RNA-seq para identificación de fusiones.....	37
Figura 8. Sobrevida global de pacientes con recaída temprana.....	43
Figura 9. Sobrevida respecto a la clasificación de riesgo NCI.....	44
Figura 10. Sobrevida respecto a variables clínicas de riesgo.....	45
Figura 11. Fusiones identificadas en LLA pediátrica y funciones biológicas asociadas.....	47
Figura 12. Identificación de subtipos moleculares con microarreglos de expresión.....	50
Figura 13. Porcentaje de subtipos genéticos identificados en LLA de linaje B.....	53
Figura 14. Fusiones del subtipo MEF2D.....	54
Figura 15. Fusiones del subtipo PAX5-r.....	56
Figura 16. Fusión P2RY8-CRLF2.....	57
Figura 17. Fusión CDKN2A-TRDC.....	58
Figura 18. Identificación de fusión CRLF2-IGH con búsqueda dirigida.....	60
Figura 19. Fusiones IGH-DUX4.....	62
Figura 20. Fusión DUX4-ETV6.....	63
Figura 21. Fusión TCF3-ZNF384.....	64
Figura 22. Fusiones sin subtipo asignado en LLA.....	66
Figura 23. Sobrevida global y sobrevida libre de evento de subtipos genéticos de LLA.....	71

ABREVIATURAS

AYA= Adolescents and Young Adults (adolescentes y adultos jóvenes)

cDNA= DNA complementario

CL= Cuenta leucocitaria

CTH= Célula troncal hematopoyética

DNA= Ácido desoxirribonucleico

EMR= Enfermedad medible residual

ETP= Early T-cell precursor (precursor temprano de células T)

FISH= Hibridación fluorescente *in situ*

GMIICLN= Grupo Mexicano Interinstitucional para la identificación de las Causas de la Leucemia en Niños

IARC= International Agency for Research on Cancer (Agencia Internacional de Investigación en Cáncer)

IC= Intervalo de confianza

IGH= Immunoglobulin heavy locus (locus de la cadena pesada de la inmunoglobulina)

ISD= Índice sociodemográfico

ITC= Inhibidores de tirosina cinasas

LALM= Leucemia aguda de linaje mixto

LLA= Leucemia linfoblástica aguda

LMA= Leucemia mieloide aguda

MO= Médula ósea

NCI= National Cancer Institute (Instituto Nacional del Cáncer, EE. UU.)

OMS= Organización Mundial de la Salud

PAR1= Región pseudoautosómica 1

PCR= Reacción en cadena de la polimerasa

RA= Riesgo alto

RH= Riesgo habitual

RNA= Ácido ribonucleico

RNA-seq= Secuenciación de RNA

RT= Recaída Temprana

RT-PCR= Reacción en cadena de la polimerasa con retrotranscripción

SG= Sobrevida global

SLE= Sobrevida libre de evento

RESUMEN

Introducción: Casi 20% de los niños mexicanos con leucemia linfoblástica aguda (LLA) fallecen en los primeros dos años a partir de su diagnóstico, y las recaídas, principalmente las tempranas, son entre 3 y 4 veces más frecuentes que en países de ingresos económicos altos. Además, un alto porcentaje (45%-70%) de casos presentan al diagnóstico características clínicas de riesgo alto (RA) de fracaso al tratamiento. La LLA puede clasificarse en subtipos moleculares con base en la presencia de alteraciones genéticas, particularmente fusiones génicas, que se asocian parcialmente con el pronóstico clínico, y su frecuencia varía con la edad y etnicidad. Menos de una cuarta parte de los casos de LLA pediátrica en México cuentan con los marcadores genéticos más comunes (*ETV6::RUNX1*, *TCF3::PBX1*, *BCR::ABL1* y *KMT2A-r*), destacando la baja frecuencia del marcador de buen pronóstico *ETV6::RUNX1*. Además, existen reportes de que fenotipos de RA y mal pronóstico como Ph-like son más comunes en poblaciones con ancestría nativoamericana. Debido a que las alteraciones de riesgo alto descubiertas recientemente son críticas para las técnicas moleculares convencionales, nuestro objetivo fue identificar fusiones génicas mediante la secuenciación del transcriptoma completo (RNA-seq) y determinar la presencia del fenotipo Ph-like en niños mexicanos con LLA.

Métodos: Realizamos un estudio de cohorte multicéntrico retrospectivo que incluyó 8 hospitales públicos de la Ciudad de México. Seleccionamos 49 pacientes para análisis de RNA-seq, y todos los pacientes fueron seguidos durante al menos 36 meses. El RNA total se extrajo de aspirados de médula ósea y la secuenciación (lecturas de extremos pareados de 76 y 151 pb) se realizó con Illumina. La detección de las fusiones génicas se realizó con el software FusionCatcher, o búsqueda dirigida basada en lecturas rasuradas (soft-clipped reads) para hallar alteraciones de *IGH*. Se usaron datos de microarreglos disponibles en las bases de datos del Laboratorio de Genómica del Cáncer para identificar casos Ph-like, y para evaluar la expresión génica de “outliers” que sugieren activación oncogénica (*CRLF2* y *DUX4*). Las fusiones génicas fueron validadas mediante RT-PCR, y secuenciación Sanger para las fusiones nuevas.

Resultados: La mediana de edad de los pacientes incluidos fue de 10.9 años, y 51% fueron del sexo femenino. Treinta y cinco (71.4%) pacientes fueron clasificados como de RA (criterios de riesgo del NCI), y 23 pacientes tuvieron recaída temprana (47%). La supervivencia global (SG) a 4 años de seguimiento fue del 50 % (IC del 95 %: 42%-71%). No se observaron diferencias en la SG y supervivencia libre de evento entre los pacientes clasificados como de RA o riesgo habitual (RH), pero los pacientes mayores de 10 años

tuvieron una SG más baja. El análisis del transcriptoma reveló que el 65.3% de los casos presentaron al menos una fusión génica. Se encontraron 30 fusiones distintas y 14 son recurrentes en LLA de linaje B. Nueve fusiones (18.4%) no se han reportado previamente, aunque alguno de los genes involucrado en la fusión está frecuentemente alterado en LLA. Agrupándolos en subtipos conocidos en LLA de linaje B, los subtipos moleculares más comunes fueron *DUX4* (13%), *CRLF2* (10,87%) y *BCR::ABL1* (10,87%). *ETV6-RUNX1* se encontró en el 6.5 %, mientras que la frecuencia de los nuevos subtipos moleculares *MEF2D*, *PAX5* y *ZNF384* fue similar a lo reportado en cohortes más grandes de otras poblaciones (6.5%, 4.3% y 2.2%, respectivamente). Como se esperaba, *TCF3::PBX1* y *ETV6::RUNX1* son marcadores de buen pronóstico y ninguno de estos pacientes falleció durante el seguimiento. EL resto de los subtipos genéticos tuvieron mal pronóstico. Los subtipos dominados por cinasas *CRLF2* y *BCR::ABL1* tuvieron una SG de 50% (7 de 10 presentaron algún evento), y los casos *P2RY8::CRLF2* tuvieron el peor desenlace (3 de 4 murieron antes de completar el tratamiento). La SG a 4 años de *DUX4* fue inesperadamente baja (33 %), en contraste con el excelente pronóstico reportado en países de alto ingreso. Todos los pacientes *DUX4* que recayeron también murieron durante el tratamiento. El grupo de pacientes sin fusiones identificadas también tuvo una baja sobrevida, de solo el 51 % a 4 años de seguimiento.

Conclusiones: A nuestro conocimiento, este es el primer análisis con RNA-seq para identificar alteraciones genéticas en LLA asociadas con mal pronóstico en México. Pudimos clasificar más del 60% de los pacientes basándonos solo en datos de transcriptoma. La implementación de estas tecnologías en un entorno clínico sigue siendo un desafío, especialmente para los países en desarrollo. Sin embargo, la identificación en este estudio de algunos subtipos moleculares con pronóstico asociado ofrece la oportunidad de implementar su detección en la clínica utilizando técnicas más económicas, explorar su valor pronóstico en cohortes más grandes, apoyar en una clasificación de riesgo más precisa y asignar tratamientos más adecuados.

Palabras clave: leucemia linfoblástica aguda, RNA-seq, secuenciación del transcriptoma, recaída temprana, fusiones génicas

ABSTRACT

Introduction: Almost 20% of Mexican children with acute lymphoblastic leukemia (ALL) die within the first 2 years from diagnosis, and relapses are 3 to 4 times more frequent than in high-income countries, mostly early relapses. Moreover, a high percentage (45%-70%) of cases have high-risk clinical features at diagnosis. ALL can be classified based on genetic subtypes that partially associate with clinical outcome, and their frequency relates with age and ethnicity. Regarding this, less than a quarter of childhood ALL cases in Mexico lack the most common genetic markers, highlighting the low frequency of the good prognosis marker *ETV6::RUNX1*. Since recently discovered, high-risk alterations are cryptic to standard molecular techniques, we aimed to identify gene fusions by whole transcriptome sequencing (RNA-seq) in a sub-cohort of Mexican children with ALL.

Methods: We carried out a retrospective multicenter cohort study that included 8 public hospitals from Mexico City. We selected 49 patients for RNA-seq analysis, and all patients were followed-up for at least 36 months. Total RNA was extracted from bone marrow aspirates and sequencing (76 and 151 bp paired-end reads) was done with Illumina. Fusion detection was performed with FusionCatcher, or targeted search based on soft-clipped reads to find *IGH* alterations. Microarray data were used to identify Ph-like cases and to assess the outlier gene expression as surrogate of oncogenic activation (*CRLF2* and *DUX4* subtypes). Known fusions were validated by RT-PCR and not previously reported ones by RT-PCR and Sanger sequencing.

Results: Thirty-five (71.4%) patients were classified as high-risk (NCI risk criteria) and 23 had early relapse (47%). The 4-year overall survival (OS) of the studied population was 50% (95% CI: 42-71). No differences in EFS and OS were observed between patients classified as high- or standard-risk based on NCI risk criteria, but patients older than 10 years had lower OS. Overall, 65.3% of patients had at least one gene fusion. Thirty different fusions were found and 14 are recurrent in B-lineage ALL. Nine fusions (18.4%) were not previously reported or had a novel partner fused to a recurrently mutated gene. By grouping them into well-known B-ALL subtypes, the most common were *DUX4* (13%), *CRLF2* (10.87%) and *BCR::ABL1* (10.87%). *ETV6::RUNX1* was found in 6.5%, in line with the low frequency reported in Latinos, whereas the frequency of newly discovered subtypes was similar to the described proportions in bigger cohorts *MEF2D*, *PAX5* and *ZNF384* (6.5%, 4.3% and 2.2%, respectively). *TCF3::PBX1* and *ETV6::RUNX1* are good prognosis markers and none of these patients died during follow-up. The remaining genetic subtypes had poor outcome. Kinase-driven subtypes *CRLF2* and *BCR::ABL1* had an OS

of 50% (7 of 10 had any event), *P2RY8::CRLF2* having the worst outcome (3 of 4 died before treatment completion). The 4-year OS of *DUX4* was unexpectedly low (33%), in marked contrast with the excellent prognosis in high-income countries. All *DUX4* patients who relapsed also died during treatment. The outcome of patients without fusions was also poor, of just 51% after 4 years.

Conclusions: To our knowledge, this is the first RNA-seq-based analysis to find genetic alterations in ALL associated with poor outcome in Mexico. We were able to classify more than 60% of patients based on transcriptome data alone. The implementation of these technologies in a clinical setting is still a challenge, especially for developing countries. Nevertheless, some molecular subtypes identified in this study offer an opportunity to implement their detection with lower-cost techniques, to explore their prognostic significance in bigger cohorts, and to provide a more precise risk classification and a more optimal treatment.

Key words: acute lymphoblastic leukemia, RNA-seq, transcriptome sequencing, early relapse, gene fusions

ANTECEDENTES

Epidemiología del cáncer infantil

La Organización Mundial de la Salud (OMS) estima que cada año son diagnosticados aproximadamente 400 mil casos de cáncer en niños y adolescentes en todo el mundo ([World Health Organization, 2021](#)). En números absolutos, el cáncer infantil es una enfermedad rara en comparación con el cáncer en adultos. Sin embargo, al ser un proceso agudo, de rápido desarrollo, y con síntomas en ocasiones poco específicos, el cáncer infantil es difícil de diagnosticar en ausencia de atención e infraestructura especializada. Lo anterior implica que cientos de miles de casos de cáncer infantil no son diagnosticados en todo el mundo (hasta más del 40%), especialmente en países de bajo y mediano ingreso económicos (BIE y MIE, respectivamente) ([Ward et al., 2019](#)).

A pesar de su baja incidencia, las consecuencias del cáncer infantil son profundas en términos personales y sociales por tratarse de una enfermedad incapacitante desde edades muy tempranas, con repercusiones en la calidad de vida a causa del tratamiento y, en muchos casos, fatal. Una métrica empleada para dimensionar el impacto en la morbilidad y mortalidad es la carga de enfermedad, que evalúa la pérdida de años de vida sana (por discapacidad o muerte) por enfermedades, accidentes y factores de riesgo a muerte ([GBD 2017 Childhood Cancer Collaborators, 2019](#)). El cáncer infantil ocupa el sexto lugar en carga de enfermedad por cáncer en cualquier edad, y casi en su totalidad es atribuible a muertes durante la infancia y adolescencia. Más del 80% de la carga mundial por cáncer infantil se encuentra en países de bajo y mediano índice sociodemográfico (ISD), países que concentran más del 90% de la población infantil en todo el mundo ([Gupta et al., 2015](#); [GBD 2017 Childhood Cancer Collaborators., 2019](#); [Stiller, 2019](#)). Esto lo hace un problema de salud pública prioritario en países BIE y MIE, en donde el diagnóstico tardío, las deficiencias en la infraestructura de salud y el tratamiento inadecuado impactan en el pronóstico de los pacientes ([Steliarova-Foucher et al., 2017](#)). Los mismos factores socioeconómicos no solo influyen en la supervivencia, sino que son la misma causa de la

subestimación de la carga real de enfermedad por las deficiencias en los registros de cáncer infantil de países en desarrollo ([Gupta et al., 2015](#); [Steliarova-Foucher, 2019](#)).

Los países de MIE atraviesan por una transición epidemiológica. En países de BIE la mayoría de las muertes por enfermedad en la niñez aún se deben a enfermedades transmisibles. En países de mediano y alto ingreso han aumentado las muertes por enfermedades crónicas en relación con la industrialización de las poblaciones ([Magrath et al., 2013](#)). Aunque no se ha demostrado una relación directa entre la industrialización y el aumento de cáncer infantil en países de mediano-alto ISD, la presencia de factores de riesgo adquiridos como la obesidad, o la exposición a ciertos agentes ambientales (benceno, radiación electromagnética, tabaco, alcohol) pudieran influir en este fenómeno ([Doll y Wakeford, 1997](#); [Pérez-Saldivar et al., 2011](#); [Zhou et al., 2014](#); [IARC Monographs on the Identification of Carcinogenic Hazards to Humans, 2020](#)). Por esta razón, el aumento en la incidencia del cáncer infantil en países con mediano-alto ISD es probablemente debido a una combinación entre cambios en el estilo de vida y exposición a factores de riesgo, pero también puede ser explicado por las mejoras en las estrategias de diagnóstico y detección de casos, especialmente en países de alto ISD ([Gupta et al., 2015](#)).

Entre 30% a 50% de las neoplasias infantiles son de origen hematológico, predominantemente leucemias. Las leucemias son un conjunto de enfermedades caracterizadas por la proliferación descontrolada de las células de origen linfóide o mielóide, las cuales suprimen la hematopoyesis normal y son capaces de invadir sitios extramedulares ([Larson, 2016](#)). En pediatría, la leucemia linfoblástica aguda (LLA) constituye más del 80% de todos los casos, seguido de las leucemias de origen mielóide (13-17%) y un porcentaje menor es de leucemias de linaje mixto (o linaje ambiguo) y leucemias crónicas (<5%) ([Robison & Bhatia, 2003](#); [Medina-Sansón, 2016](#)). De la LLA, más del 80% de los casos de LLA poseen el inmunofenotipo de precursores/progenitores B (LLA-preB), y el resto de linaje T o fenotipos mixtos (bifenotípicas) ([Medina-Sansón, 2016](#)).

La LLA surge a partir de la transformación maligna de un precursor de estirpe linfoide (linaje B o T) como resultado de la acumulación de mutaciones somáticas que impiden su maduración y le confieren otras propiedades como la resistencia a la apoptosis y desregulación del ciclo celular (Larson, 2016).

El pronóstico de los pacientes con LLA ha mejorado sustancialmente a lo largo de las últimas décadas en países desarrollados, en donde hasta el 98% de los casos alcanzan la remisión clínica durante las primeras semanas de tratamiento. Esto no ocurre en países en desarrollo, en donde además del aumento en la incidencia de la enfermedad, las tasas de recaída y mortalidad siguen siendo muy altas. Los factores socioeconómicos son determinantes para el éxito del tratamiento, pero también existen evidencias de que algunos factores biológicos pueden explicar las notables diferencias en la incidencia y mortalidad de LLA entre las poblaciones. Esto pone especial énfasis en las regiones con menos recursos, en donde viven la mayor cantidad de niños de todo el mundo (Magrath et al., 2013).

Etiología y origen de la LLA

La mayoría de los casos de LLA no tienen una causa conocida, y solo una minoría se asocian a síndromes genéticos que predisponen al desarrollo de diversas neoplasias hematológicas. Es necesaria la interacción entre distintos factores como exposición a agentes ambientales, predisposición genética, y su interacción en una ventana de tiempo de vulnerabilidad que hace posible su desarrollo (Inaba et al., 2013; Mejía-Aranguré, 2016). Algunos procesos mutagénicos son intrínsecos a la ontogenia linfoide y son ineludibles (Machado et al., 2022). Otros factores de riesgo pueden ser modificables, como la obesidad y su posible asociación para desarrollar LLA o presentar un peor pronóstico a la enfermedad (Larsson & Wolk, 2008; Gelelete et al., 2011; Hjalgrim et al., 2013; Orgel et al., 2014; Jiménez-Hernández et al., 2018; Núñez-Enríquez et al., 2019). La radiación ionizante sigue siendo el único agente causal en LLA reconocido hasta ahora, y el resto de los agentes ambientales carecen de suficiente evidencia o plausibilidad biológica (Preston et al., 1994; Bartley et al., 2010; Coglianò et al., 2011; Inaba et al., 2013).

Por lo tanto, se considera que en la gran mayoría de los casos de LLA y tumores pediátricos en general, la alteración genética fundadora es un defecto del desarrollo surgido durante etapas embrionarias (Ma et al., 2018). Esta idea está apoyada por la ausencia de una etiología clara, una menor latencia postnatal para su aparición clínica, y una baja carga mutacional en comparación con los tumores en edad adulta (Greaves, 2018; Gröbner et al., 2018; Ma et al., 2018).

Las hipótesis más aceptadas sugieren que la alteración genética fundadora que da origen a la LLA pediátrica surge *in utero* en una célula troncal hematopoyética (CTH) o en un progenitor linfoide, capaz de generar una población de células pre-leucémicas que requieren de más alteraciones cooperadoras para la aparición de la enfermedad (Greaves, 2003; Greaves, 2006). Sin la presencia de eventos o factores de riesgo que promuevan la adquisición de mutaciones secundarias, la leucemia no se desarrollaría. En LLA, una respuesta aberrante a infecciones comunes se ha sugerido como un evento que promueve la adquisición de mutaciones cooperadoras para la transformación maligna. De acuerdo con esto, existen dos hipótesis sobre el surgimiento de la LLA. La hipótesis de Kinlen postula que una interacción inusual entre individuos susceptibles con individuos contagiados con algún agente infeccioso común induce una respuesta inmunológica excesiva (Kinlen, 1998). Por otra parte, la hipótesis de Greaves sugiere que la LLA surge por una exposición tardía a infecciones comunes en individuos susceptibles, lo que ocasionaría una respuesta aberrante en un sistema inmune previamente no estimulado (Greaves, 1988; Greaves, 2018). Esta respuesta incontrolada, considerando la capacidad fisiológica de recombinación somática de las células linfoides, podrían generar alteraciones genéticas en clonas pre-leucémicas en individuos con cierto grado de susceptibilidad (Inaba et al., 2013). Esto se consideraría una consecuencia de la modernización y del cambio a un estilo de vida más “higiénico”, que al mismo tiempo buscaba explicar por qué la incidencia de esta enfermedad parecía incrementarse en los países desarrollados (Greaves, 2018). Estas hipótesis también explicarían por qué solo algunos

individuos desarrollan LLA, ya que la presencia de clonas pre-leucémicas es un fenómeno no solo más frecuente de lo esperado, sino incluso, común.

Desarrollo prenatal de la LLA

A nivel molecular, la LLA puede considerarse como un conjunto de enfermedades definidas por la presencia de una gran diversidad de alteraciones genéticas, cada una con características biológicas distintas. Su curso clínico y pronóstico también se asocia en cierta medida a la presencia de estas alteraciones. Existe evidencia sólida de que, al menos para algunos de los subtipos más comunes de LLA-B, el primer “hit” o alteración genética surge antes del nacimiento (Figura 1). Esto se ha descrito también para otros tipos de cáncer pediátrico, aunque no para la LLA de linaje T (LLA-T), la cual es más frecuente en la adolescencia (Greaves & Wiemels, 2003; Marshall, 2014).

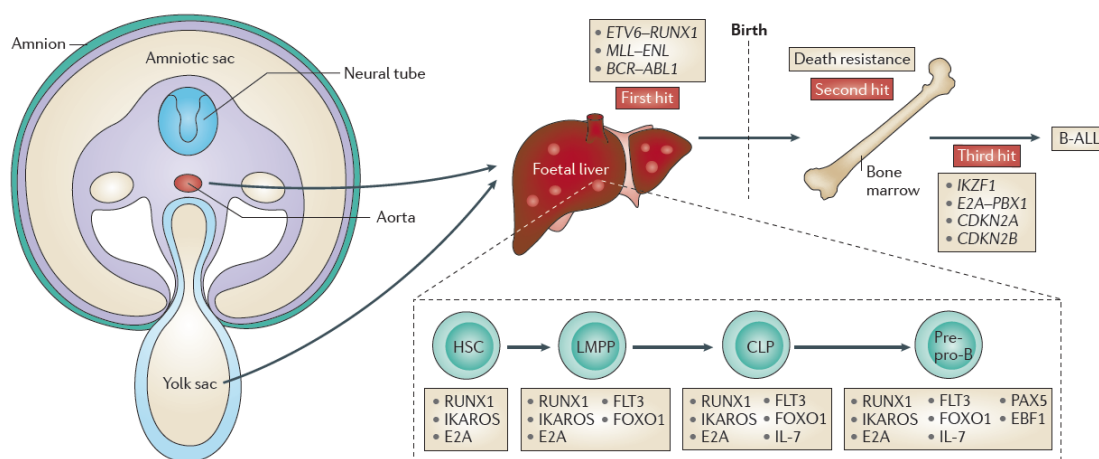


Figura 1. Desarrollo de la LLA de linaje B. La presencia de una fusión en desde etapas pre-natales surge en una célula troncal hematopoyética o en un progenitor linfoide en alguna de sus distintas etapas de maduración. El primer “hit” ocurre durante la hematopoyesis temprana, la cual inicia en el endotelio homogénico del saco vitelino y en la región aorta-gonadal-mesonefros, se traslada al hígado fetal hasta finalmente alojarse en la médula ósea en el período perinatal en adelante. La primera alteración suele ser una fusión génica, en el que se involucran genes indispensables para la maduración de las hematogonias. Las clonas pre-leucémicas necesitan de subsecuentes “hits” después del nacimiento para el desarrollo de la enfermedad. Tomado y modificado de Marshall et al., 2014. HSC: Hematopoietic stem cell; LMPP: Lymphoid-primed multipotent progenitors; CLP: Common lymphoid progenitor; Pre-pro-B: Precursor-progenitor-B.

Dos líneas de evidencia apoyan las hipótesis del origen prenatal de la LLA: concordancia de su desarrollo en gemelos idénticos, y estudios en sangre de neonatos y la identificación de alteraciones genéticas fundadoras desde el nacimiento. El primer reporte en la literatura de una leucemia aguda concordante en gemelos idénticos data de 1882, seguido de varios reportes de casos publicados a finales de la década de los 50's ([Senator H, 1882](#); [Wolman 1962](#)). La concordancia temporal en el desarrollo de LLA entre gemelos idénticos para el subtipo con alteraciones del gen *KMT2A* (antes *MLL*), un subtipo muy agresivo y predominante en niños menores de 1 año, es prácticamente del 100%. Ya que la mayoría de los gemelos idénticos (o monocigóticos) comparten la misma placenta (monocoriónicos), se propuso que la leucemia pudiera haber iniciado en uno de los individuos, y fue diseminada al otro vía anastomosis placentaria ([Greaves & Hughes, 2018](#)). Ahora se sabe que el origen monoclonal de la LLA en gemelos idénticos ocurre de forma similar para otros subtipos como *ETV6::RUNX1*, aunque para este su latencia puede ser prolongada (pueden pasar varios años de diferencia en ser diagnosticados) y menos concordante (~10-15%) ([Ford et al., 1998](#); [Wiemels et al., 1999](#); [Teuffel et al., 2004](#)). En gemelos idénticos, la probabilidad para desarrollar leucemia no es distinta al resto de la población infantil, pero sí es muy superior para el otro individuo cuando uno de ellos la desarrolla. Su excepcional sincronismo y alta concordancia, esto para el subtipo *KMT2A*, hacía difícil pensar en una susceptibilidad genética compartida. Sin embargo, no se podía descartar la exposición *in utero* a algún agente genotóxico no identificado. Aun así, la predisposición genética o la exposición a algún agente ambiental eran ideas más aceptadas en un inicio que la diseminación por la placenta ([Greaves et al., 2003](#)). Fue hasta la identificación de marcadores genéticos distintivos de las leucemias agudas con el desarrollo de nuevas herramientas tecnológicas, que se pudo comprobar la hipótesis de un origen clonal único en gemelos y que el primer hit en la leucemogénesis puede ocurrir en individuos que no llegan a desarrollar LLA. Retrospectivamente se pudo obtener y analizar sangre de neonatos mediante tarjetas Guthrie (trozos de papel filtro usadas para coleccionar gotas de sangre de neonatos), usadas para identificar trastornos

metabólicos al nacimiento. La fusión *KMT2A::AFF1* (anteriormente llamada *MLL::AF4*) fue identificada en pacientes que desarrollaron LLA al paso de algunos meses o durante los primeros años, confirmando su surgimiento prenatal (Gale et al., 1997). Esto reforzó la idea de que las fusiones del gen *KMT2A*, por su rápida aparición y alta concordancia en gemelos idénticos, son suficientes o requieren pocas mutaciones adicionales para el desarrollo de la enfermedad (Ford et al., 1993; Andersson et al., 2015). Sin embargo, esta alteración es muy poco frecuente, se presenta casi de forma exclusiva en niños menores de un año (incluso congénita), y no coincide con el primer y mayor pico de incidencia de la LLA (entre los 2-5 años). La siguiente interrogante era saber qué explicaba el surgimiento de las alteraciones más comunes en LLA. También se descubrió con tarjetas Guthrie que otras alteraciones más comunes en LLA no solo surgen también *in utero*, son de hecho frecuentes, y exceden en órdenes de magnitud la incidencia de la enfermedad (Taub et al., 2002). La fusión *ETV6::RUNX1* es uno de los marcadores más frecuentes en LLA pediátrica y con un patrón de incidencia sincrónica al de la propia enfermedad. Se pudo comprobar que esta alteración estaba presente desde el nacimiento, consistente con su origen prenatal tanto en casos individuales como en gemelos idénticos (Wiemels et al., 1999). A diferencia de las fusiones con *KMT2A*, la evidencia apuntaba a que *ETV6::RUNX1* era tal vez la alteración inicial pero insuficiente por sí misma para causar la leucemia. Estos hallazgos son la base del modelo que ahora conocemos como “doble hit” (diferente al modelo de Knudson de inactivación de supresores tumorales) de la leucemogénesis. Según este modelo, se requiere la presencia de una clona o una población clonal de células pre-leucémicas *in utero* con una alteración genética fundadora, y la ganancia de mutaciones cooperadoras después del nacimiento para la progresión y aparición clínica de la enfermedad (McHale et al., 2003; Bateman et al., 2010; Teuffel et al., 2004; Ma et al., 2013). Cuando estas alteraciones se buscaron en grandes cohortes de tarjetas Guthrie o en bancos de sangre de cordón umbilical (los ensayos en tarjetas Guthrie pueden tener baja sensibilidad o un número insuficiente de clones detectables) se estimó que hasta 1-2 de cada 100 nacimientos pueden albergar clones *ETV6::RUNX1+* (Mori et al., 2002; Zuna et al., 2011). Esta

frecuencia ha sido cuestionada en otros trabajos ([Lausten-Thomsen et al., 2011](#)), aunque otras fusiones como *TCF3::PBX1* y *P2RY8::CRLF2* también se han identificado en sangre neonatal ([Morak et al., 2012](#); [Hein et al., 2019](#)).

Aunque es aún controversial, múltiples evidencias respaldan la hipótesis del desarrollo prenatal de la LLA, y la necesidad de subsecuentes “hits” para la manifestación de la enfermedad. La incógnita sigue siendo si este fenómeno ocurre para todos los subtipos de LLA. Las evidencias hasta ahora solo respaldan este modelo para las alteraciones típicas en el primer y mayor pico de incidencia de la enfermedad, en especial las fusiones *ETV6::RUNX1*, *TCF3::PBX1* y las hiperdiploidías. Aunque muy frecuentes antes de los 6 años, estos subtipos son infrecuentes en edades más avanzadas.

Nuevo panorama genético de la LLA

La LLA se puede clasificar por la presencia de alteraciones genéticas que definen subtipos con un comportamiento clínico y biológico únicos (Figura 2). Como ocurre en otros tumores pediátricos, en esta leucemia predominan las aberraciones en la estructura (traslocaciones, deleciones) o el número (aneuploidías) de cromosomas, mientras que las mutaciones puntuales son poco frecuentes. Estas alteraciones son características de la enfermedad y su linaje, y a menudo los genes afectados son esenciales para la maduración de los linfocitos. En LLA-B, la mitad de todas las alteraciones genéticas son identificables mediante citogenética convencional, cariotipo, e hibridación por fluorescencia *in situ* (FISH) ([Inaba y Pui, 2021](#)). Las más comunes surgen como resultado de traslocaciones cromosómicas que dan origen a fusiones génicas y proteínas quiméricas con actividades ajenas a su función normal, como las fusiones *ETV6::RUNX1*, *TCF3::PBX1*, *BCR::ABL1*, y los rearrreglos del gen *KMT2A* (Figura 2). Otras involucran la ganancia (hiperdiploidías) o pérdida (hipodiploidías) de cromosomas, siendo la hiperdiploidía la alteración más común en LLA-B pediátrica. El panorama para la LLA-T es menos claro,

con algunas alteraciones distintivas para este linaje, pero sin la misma asociación pronóstica que las alteraciones en LLA-B.

Fue a partir del surgimiento de herramientas de análisis masivo como los microarreglos y la secuenciación que se pudo caracterizar un importante porcentaje de casos de LLA en los que no se identificaba ninguno de los marcadores genéticos más conocidos. Estos nuevos subtipos también pudieron asociarse con características clínicas y relevancia pronóstica (la mayoría de ellos presentes en pacientes de alto riesgo) e identificar oportunidades terapéuticas que están siendo incorporadas a los tratamientos convencionales.

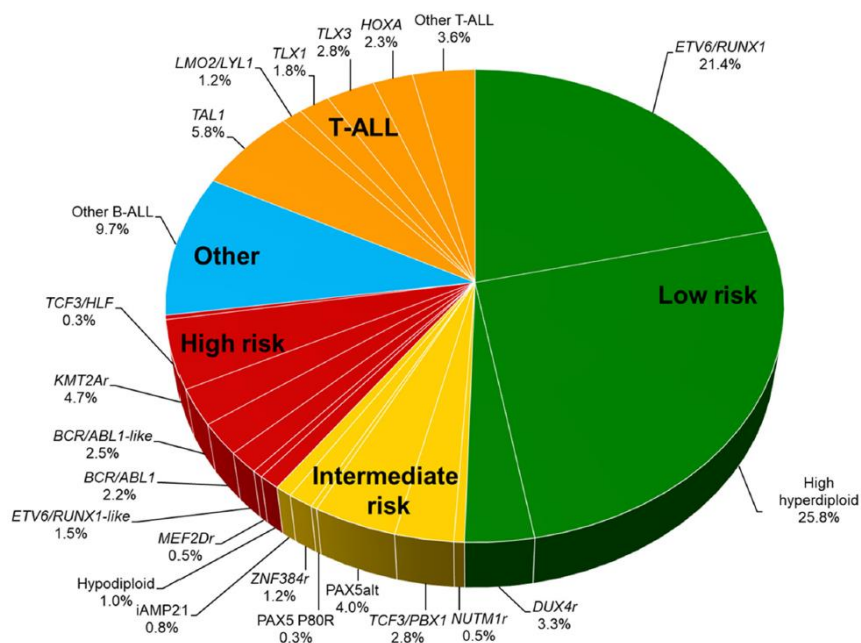


Figura 2. Subtipos genéticos en LLA. Los porcentajes estiman la prevalencia aproximada en LLA infantil de la iniciativa TARGET (Therapeutically Applicable Research to Generate Effective Treatments) en Estados Unidos. Los subtipos de linaje B tienen un pronóstico asociado, mientras que en la LLA de linaje T no hay una clasificación genética con pronóstico claro asociado. Tomado de Inaba y Pui, 2021.

Otros B

El término "Otros B" (del inglés "B-others") define un conjunto de subtipos con alteraciones genéticas distintas a los marcadores comunes en LLA (aneuploidías, *ETV6::RUNX1*, *TCF3::PBX1*, *BCR::ABL1*, *KMT2A-*

r). Fue hasta el desarrollo de tecnologías con mayor poder de resolución y el análisis de grandes cohortes de pacientes que se identificaron como nuevos subtipos en LLA, aunque algunos ya habían sido descritos en reportes de caso. Algunas de estas alteraciones suelen ser indetectables con técnicas moleculares convencionales (crípticas).

La presencia de múltiples fusiones génicas entre un gen conductor común con múltiples genes compañeros (del inglés “partner genes”) también dificultaba su identificación. El ejemplo conocido es *KMT2A*, del cual se conocen más de 100 fusiones que lo involucran (Meyer et al., 2018). Actualmente se sabe que esto ocurre en fusiones que involucran genes como *MEF2D* o *PAX5* en posición 5', y *ZNF384* en posición 3'. En algunas ocasiones la activación oncogénica surge de la translocación de un proto-oncogén a promotores altamente activos o potenciadores, denominado secuestro de potenciadores (“enhancer hijacking”), y no necesariamente resulta en la generación de una proteína quimérica. Este es el caso de las alteraciones de genes como *DUX4*, *CRLF2*, y *CEBPE*. Algunos subtipos fueron identificados con perfiles de expresión génica global, primero con microarreglos y después con secuenciación de RNA. Por ejemplo, se descubrió que algunos pacientes presentaban perfiles de expresión génica similares al de alteraciones genéticas comunes, pero sin la presencia de estas. El caso más conocido es *BCR::ABL1*-like (o Ph-like), y recientemente *ETV6::RUNX1*-like o *ZNF384*-like.

Actualmente se sabe que “Otros B” representan entre 20% y 30% de los casos de LLA sin alteraciones recurrentes. Al igual que los marcadores clásicos, tienen además características y pronóstico clínico asociados e incluso pueden ser blanco de terapias dirigidas (Lilljebjörn et al., 2017).

Alteraciones del gen *DUX4*.

El subtipo molecular *DUX4* fue identificado en uno de los primeros estudios de microarreglos de expresión en LLA, pero la alteración conductora subyacente no pudo ser detectada hasta que se realizó secuenciación de transcriptoma (Yeoh et al., 2002). Primero se reportó en un grupo de pacientes que

presentaban deleciones frecuentes en el gen *ERG* y con un perfil de expresión particular (Harvey et al., 2010). Con RNA-seq se identificó que la alteración distintiva de este nuevo subtipo es la activación del gen doble homeobox 4 (*DUX4*), expresado casi exclusivamente en etapas embrionarias tempranas y en testículos (De Iaco et al., 2017; Li et al., 2018). *DUX4* se localiza en la región macrosatélite subtelomérica D4Z4 del brazo largo del cromosoma 4, y una altamente homóloga en el cromosoma 10 (Rehn et al., 2020). Cada repetido D4Z4 (10-100 repetidos por cromosoma) contiene un marco de lectura para el gen *DUX4*, el cual carece de intrones y codifica para una proteína con dos homeodominios (Dong et al., 2018). Es paradójico que su activación está involucrada en procesos pro-apoptóticos en tejidos somáticos, por lo que cuando se expresa en células musculares causa distrofia muscular facioescapulohumeral, una enfermedad degenerativa (Choi et al., 2016). La yuxtaposición de un segmento de D4Z4 a secuencias promotoras o potenciadoras (como la fusión *CIC-DUX4*) también ha sido identificada en sarcomas pediátricos (Yoshimoto et al., 2017). En LLA la alteración más común transloca a *DUX4* a las secuencias potenciadoras o promotores del locus de la cadena pesada de las inmunoglobulinas (*IGH*), situado también en la región subtelomérica del cromosoma 14 (Yasuda et al., 2016). La expresión inducida de *DUX4::IGH* detiene la diferenciación de precursores linfoides y genera LLA-B *in vivo* (Yasuda et al., 2016; Tanaka et al., 2018). Los blancos transcripcionales de *DUX4::IGH* no son los mismos que la versión wild-type de *DUX4*, en parte porque la mayoría de alteraciones de *DUX4* generan un extremo c-terminal trunco, crítico para la transactivación de sus genes blanco (Tanaka et al., 2018). En el caso de las alteraciones con *IGH*, la translocación al alelo silenciado de *IGH* podría atenuar los efectos tóxicos que resultarían de la excesiva expresión de *DUX4* como ocurre con su versión wild-type, los cuales son pro-apoptóticos (Tian et al., 2019). Se ha propuesto que la capacidad de transformación de *DUX4* se debe a que activa la expresión de una forma alternativa de *ERG* y la pérdida de su función, esencial para la hematopoyesis definitiva (Zhang et al., 2016; Dong et al., 2018). Sin embargo, aún no se conocen muchos de los mecanismos por los cuales las proteínas alteradas *DUX4*

inducen el desarrollo de leucemia. *DUX4-r* está presentes en aproximadamente 5%-7% de los casos de LLA-B, aunque es más frecuente en adolescentes y adultos jóvenes y en población asiática (Lilljebjörn et al., 2016; Yasuda et al., 2016; Li et al., 2021). A pesar de la presencia de marcadores mieloides y persistencia de enfermedad medible residual (EMR) post-inducción, el pronóstico de este subtipo es excelente. Las tasas de supervivencia global de *DUX4* son superiores al 95% y 85% en niños y AYAs (del inglés Adolescents and Young Adults), respectivamente (Yeoh et al., 2002; Schinnerl et al., 2019; Li et al., 2021; Zhang et al., 2016).

Rearreglos del gen *ZNF384* (*ZNF384-r*).

Las fusiones que involucran a la proteína dedo de zinc 384 (*ZNF384*) representan un grupo de leucemias agudas de linaje ambiguo, clasificadas como LLA, pero con expresión de marcadores mieloides o de linfoblastos T. Las primeras alteraciones con *ZNF384* (gen en posición 3') se identificaron en LMA y leucemia aguda de linaje mixto (LALM) por la presencia aberrante de marcadores mieloides como CD13 o CD33 (Martini et al., 2002; La Starza et al., 2005). *ZNF384* codifica para un factor de transcripción que regula la expresión de genes involucrados en la remodelación de la matriz extracelular (Bidwell et al., 2001). Las proteínas de fusión en este grupo conservan los motivos zf-C2H2 de unión a DNA del extremo carboxilo terminal de *ZNF384*, y diferentes proteínas en el extremo amino terminal como TCF3, EP300, TAF15 y EWSR1 entre las más comunes (Zhong et al., 2008; Nyquist et al., 2011; Gocho et al., 2015). Las alteraciones *ZNF384-r* se encuentran entre el 2% al 5% de los casos de LLA pediátrica, y son más frecuentes en AYAs (Shago et al., 2016; Hirabayashi et al., 2021). *ZNF384-r* manifiesta un perfil transcripcional similar al de células troncales hematopoyéticas, activación de factores de transcripción de linaje mieloides como *GATA3* y activación de vía JAK-STAT (Yaguchi et al., 2017; McClure et al., 2018; Dickerson et al., 2022). Uno de los genes con mayor expresión en este subtipo es el gen *FLT3* (tirosina cinasa similar a *fms 3*), y el uso de inhibidores de esta cinasa en estudios preclínicos ha mostrado eficacia

como posible terapia dirigida (Alexander et al., 2018; Dickerson et al., 2022). Aunque los casos *ZNF384-r* suelen presentar una mala respuesta inicial al tratamiento y frecuente cambio de linaje con o sin tratamiento, se considera de pronóstico intermedio si se trata con protocolos convencionales de quimioterapia intensiva (Hirabayashi et al., 2017; Novakova et al., 2021).

Rearreglos del gen *MEF2D* (*MEF2D-r*).

El gen *MEF2D* codifica para el factor potenciador específico de miocitos 2D, un activador transcripcional involucrado en procesos de crecimiento, supervivencia celular y apoptosis en células musculares y neuronas (Potthoff y Olson, 2007). *MEF2D*, junto con otros miembros de la familia MEF, interactúan con proteínas involucradas en procesos de acetilación y desacetilación de histonas (Lu et al., 2000; Youn et al., 2000; Youn et al., 2000). La inactivación de *MEF2D* detiene la diferenciación de precursores linfoides en fase pre-B (Herglotz et al., 2016). El primer rearreglo identificado con este gen fue en una línea celular pre-B con una translocación entre los cromosomas 1 y 19 (que también dan origen a la fusión recurrente *TCF3::PBX1*), resultando en la fusión *MEF2D::DAZAP1* (Yuki et al., 2004). La proteína quimérica resultante tenía una mayor actividad transcripcional de los genes blanco de *MEF2D* que el de la proteína en estado silvestre (Prima et al., 2005). Todas las fusiones que involucran a *MEF2D* conservan su extremo amino terminal, el cual contiene el dominio MADS de unión a DNA (Shore & Sharrocks, 1995; Wu et al., 2011). Se ha sugerido que la proteína del extremo carboxilo terminal estabiliza e incrementa la actividad transcripcional de *MEF2D*, razón por la cual se observa el mismo perfil de expresión en casi todos los casos *MEF2D-r* (con excepción de *MEF2D::CSF1R* clasificado como Ph-like) (Gu et al., 2016). Este subtipo posee un inmunofenotipo característico (CD38+ y ausencia o baja expresión de CD10), similar al de células B maduras, predominantemente en pacientes de mayor edad y asociado a un mal pronóstico clínico (Suzuki et al., 2016). Se estima que *MEF2D-r* representa alrededor de 3-5% de los casos de LLA-B y es más común en AYAs (Liu et al., 2016; Gu et al., 2016). Actualmente se conocen 8 genes de fusión que

involucran a *MEF2D* en LLA-B, entre los que se encuentran *BCL9*, *SS18*, *HNRNPUL1* (Gu et al., 2016; Li et al., 2018). Algunas potenciales vulnerabilidades que se han identificado para este subtipo están los inhibidores de desacetilasas de histonas (HDACs), o el uso de inhibidores de cinasas SYK y SRC que bloquean la señalización pre-BCR, crucial para la sobrevivencia de blastos leucémicos con fusiones de *MEF2D* (Tsuzuki et al., 2020).

Fenocopias de subtipos recurrentes, Ph-like y *ETV6::RUNX1*-like.

En 2009 se identificó a un conjunto de pacientes que presentaban un perfil de expresión génica similar a *BCR::ABL1* (o cromosoma Philadelphia, Ph) pero sin la presencia de esta fusión canónica (Den Boer et al., 2009; Mullighan et al., 2009). Inicialmente la alteración genética conductora no estaba identificada, pero los casos *BCR::ABL1*-like (o Ph-like) tenían deleciones frecuentes en genes esenciales para la maduración de linfocitos B (como *IKZF1* o *PAX5*) como los *BCR::ABL1*+ (Den Boer et al., 2009; Mullighan et al., 2009). Ph-like está compuesto por distintos subtipos definidos por alteraciones en genes que codifican para proteínas con actividad cinasa y/o receptores de citocinas. El subtipo más común de Ph-like se caracteriza por la sobre-activación del receptor de la linfopoyetina tímica derivada del estroma (TSLPR), codificado por el gen *CRLF2* (Russell et al., 2009). Dos alteraciones son las más comunes, ambas conducen a la sobre-expresión de *CRLF2: P2RY8::CRLF2*, surgida de una deleción críptica de ~320 kb en la región pseudoautosómica (PAR1) de los cromosomas sexuales, e *IGH::CRLF2*, producida por la translocación de la secuencia entera de *CRLF2* al locus *IGH* (Russell et al., 2009; Harvey et al., 2010). Los rearrreglos génicos de *CRLF2* (*CRLF2-r*) representan aproximadamente la mitad de todos los casos Ph-like, tienen mutaciones puntuales frecuentes en genes de cinasas JAK, y son la alteración más frecuente en pacientes con LLA y síndrome de Down (Russell et al., 2009; Mullighan et al., 2009; Harvey et al., 2010). El segundo subtipo más frecuente, denominado “clase ABL”, se caracteriza por presentar fusiones entre genes como *ABL1/2*, *PDGFRB* y *CSF1R* en posición 5', y múltiples genes compañeros en posición 3'

(Roberts et al., 2014, Tasian et al., 2017). Otros genes involucrados con menor frecuencia en Ph-like están *EPOR*, *JAK2*, *TYK2*, *NTRK3* y *FLT3* (Roberts et al., 2012; Roberts et al., 2014). La gran cantidad de rearrreglos génicos de Ph-like identificados se conoció hasta el empleo de secuenciación masiva, pero la determinación de Ph-like se realiza con perfiles de expresión de paneles de genes o expresión global, con algunas discordancias dependiendo de la herramienta empleada (Boer et al., 2015; Reshmi et al., 2017). No obstante, es la alteración genética subyacente la que es clínicamente accionable. Al igual que *BCR::ABL1*, el uso de inhibidores de tirosina cinasa (ITC) puede incorporarse a los esquemas terapéuticos de Ph-like, como el ruxolitinib para *CRLF2-r*, rearrreglos de *EPOR* y *JAK2*, e imabinitib/dasatinib para el tratamiento de los casos de clase ABL (Figura 3). Como ocurre en los subtipos "Otros-B", la frecuencia de Ph-like varía en función de la edad, el grupo de riesgo y la etnicidad. Ph-like representa desde ~5% -10% de los casos con ascendencia asiática, africana, o clasificados como RH, y más del 15% en niños con características de RA y con ascendencia nativo-americana/latina (Harvey et al., 2010; Roberts et al., 2014; Li et al., 2018; Lee et al., 2022). Su pronóstico clínico es excepcionalmente pobre y con tendencia a mala respuesta al tratamiento convencional, lo que resalta el potencial de los ITC los cuales ya están siendo incorporados como terapia adyuvante en ensayos clínicos formales (Maese et al., 2017).

Otra fenocopia de un subtipo común sin la presencia de la alteración canónica es *ETV6::RUNX1*-like, que representa entre el 1% al 5% de los casos de LLA (Lilljebjörn et al., 2017). En contraste con Ph-like y Otros B, es progresivamente menos frecuente en AYAs y adultos (Gu et al., 2019). En *ETV6::RUNX1*-like predominan las fusiones de *ETV6* y alteraciones en *IKZF1*, aunque estas últimas parecen no conferir un peor pronóstico como ocurre en *BCR::ABL1+* y Ph-like (Lilljebjörn et al., 2016). Debido a su baja frecuencia, no es claro su pronóstico clínico. Sin embargo, un mayor porcentaje de estos casos suelen ser positivos a EMR post-inducción, en contraste con *ETV6::RUNX1* cuya respuesta al tratamiento es excelente (Li et al., 2018; Gu et al., 2019). La OMS designó a Ph-like y *ETV6::RUNX1*-like como entidades únicas en linfoma/leucemia linfoblástica de linaje B en su última clasificación de tumores hemato-

linfoides (Alaggio et al., 2022). Un subtipo recién descrito es *ZNF384*-like. Este subtipo no cuenta con fusiones de tipo *ZNF384*, pero el truncamiento de su extremo carboxilo terminal le confiere características moleculares e inmunofenotípicas idénticas al subtipo con la fusión canónica (Zaliova et al., 2021).

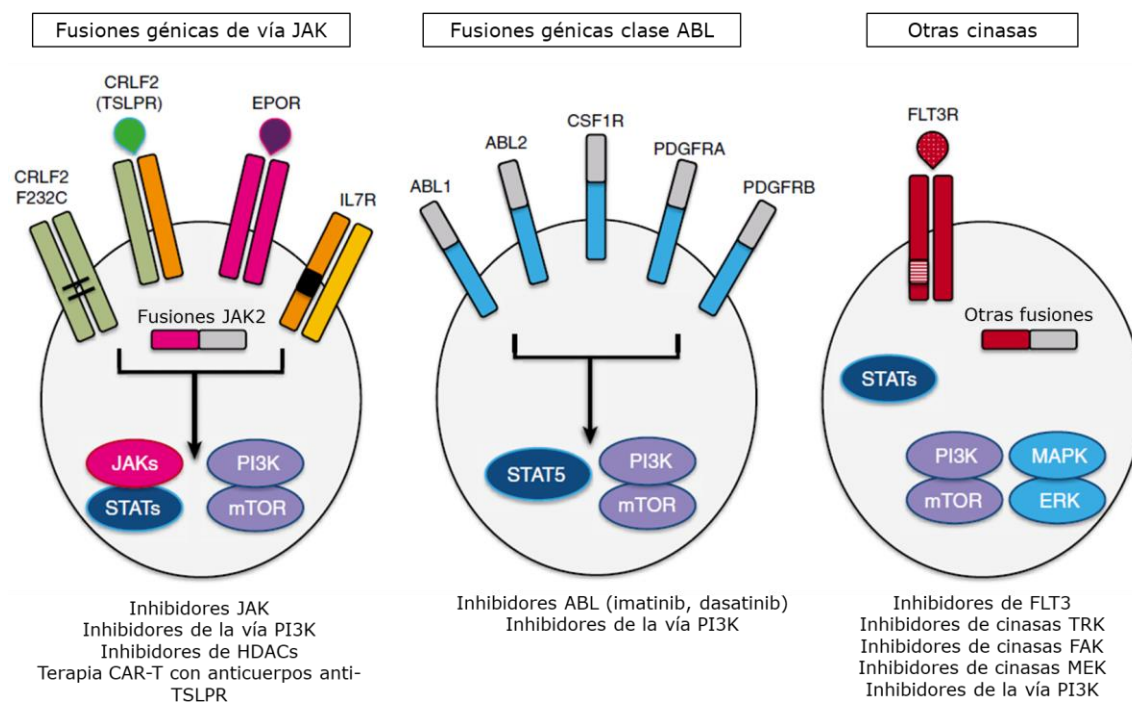


Figura 3. Cinasas activadas en LLA Ph-like. Fusiones génicas y otras alteraciones en cinasas y receptores de citocinas potencialmente accionables con inhibidores tirosina cinasa. La clasificación está basada en mayor medida a la sensibilidad acorde al tipo de cinasa o vía de señalización activa. Tomado y modificado de Tasian et al., 2017. HDACs: Inhibidores de desacetilasas de histonas.

Alteraciones conductoras de *PAX5*: *PAX5alt* y *PAX5-P80R*.

El gen Paired Box 5 *PAX5* (anteriormente BSAP, activador específico de linaje de células B) es un factor de transcripción crítico para la diferenciación de los linfocitos B (Cobaleda et al., 2007; Okuyama et al., 2019). Las alteraciones somáticas de *PAX5* están presentes en más de 30% de los casos de LLA de linaje B, consideradas como lesiones cooperadoras importantes para el arresto en la diferenciación de

linfocitos B (Mullighan et al., 2007). Sin embargo, un grupo de pacientes con presencia de fusiones de *PAX5* (en extremo 5') representan un subtipo único y poseen un perfil de expresión génica que en ocasiones se traslapa con casos *CRLF2-r/Ph-like* (Li et al., 2018). La excepción es *PAX5::JAK2*, una fusión que resulta en la activación de la cinasa JAK2 y se agrupa como Ph-like (Roberts et al., 2014; Schinnerl et al., 2015; Lilljebjörn et al., 2016; Gu et al., 2019). Aún no es claro por qué algunas fusiones de *PAX5* actúan como alteraciones conductoras o cooperadoras, pero casi siempre es necesaria la pérdida del alelo no mutado de *PAX5* (Jurado et al., 2022). *PAX5*-altered (*PAX5alt*) se identifica principalmente con perfiles de expresión génica y el ~40% presentan fusiones con *PAX5*, aunque también alrededor de 30% presentan otras alteraciones como mutaciones puntuales y duplicaciones intragénicas (Gu et al., 2019; Zaliova et al., 2019). *PAX5alt* está presente entre el 3% de casos de riesgo habitual (RH) al ~10% de LLA-B y en niños con características de riesgo alto (RA) y adolescentes, aunque se considera de pronóstico intermedio (Gu et al., 2019). Solo una mutación puntual en *PAX5* representa un subtipo distinto a *PAX5alt*, en la que ocurre un cambio de una prolina por una arginina en el residuo 80 dentro del dominio de unión a DNA (*PAX5* p.Pro80Arg o *PAX5* P80R) (Mullighan et al., 2007; Li et al., 2018). Los casos de *PAX5* P80R poseen un perfil de expresión único y tienen un efecto deletéreo mayor en la diferenciación hacia linaje B que *PAX5alt* (Gu et al., 2019; Li et al., 2018). *PAX5* P80R es poco frecuente en niños (~1) y más frecuente en AYA y adultos (~5%) (Roberts et al., 2018; Zaliova et al., 2019; Lee et al., 2022). El pronóstico de *PAX5* P80R difiere considerablemente entre cohortes analizadas, pero suele ser peor que *PAX5alt* (Passet et al., 2019; Gu et al., 2019; Jung et al., 2020). Al igual que *DUX4-r* y *ZNF384-r* exhiben plasticidad de linaje antes o después del tratamiento, lo cual tiene importantes implicaciones en el monitoreo de blastos residuales y para el éxito de los tratamientos inmunológicos dirigidos contra moléculas específicas de linaje B (Novakova et al., 2021; Wu et al., 2021).

Otras alteraciones del locus IGH.

Además de *CRLF2* y *DUX4*, otros oncogenes se activan cuando son translocados al locus *IGH*, aunque son subtipos/alteraciones poco frecuentes presentes en menos del 1% de los casos. Algunas están incluidas dentro de subtipos comunes, como *IGH::EPOR*, recurrente en Ph-like (Iacobucci et al., 2016). Otras constituyen subtipos únicos, como *IGH::CEBPE*, que conduce a la activación de la proteína ϵ de unión a CCAAT/enhancer CEBPE, un factor de transcripción esencial para la diferenciación de granulocitos (Pierini et al., 2011; Zerrouki et al., 2016; Shyamsunder et al., 2019). *IGH::CEBPE* posee un perfil transcripcional único mutaciones concomitantes frecuentes en el gen *ZEB2* (Li et al., 2018; Lilljebjörn et al., 2016). Ocasionalmente se presenta de forma concomitante con otros subtipos, por lo que su pronóstico y variables clínicas asociadas no son claras (Gu et al., 2019). La activación de oncogenes comunes en cáncer como *MYC* y *BCL2* también constituyen un subtipo único pero presente en una minoría de casos en LLA, similares a las descritas en linfomas de “doble hit” (Zhuang et al., 2022). Las alteraciones *IGH::MYC* e *IGH::BCL2* están presentes casi de forma exclusiva en AYAs y adultos y tienen un pronóstico excepcionalmente pobre (Russell et al., 2014; Gu et al., 2019).

Nuevos subtipos de LLA-T

Análisis genómicos han permitido identificar subtipos de LLA de linaje T con comportamiento biológico particular. No obstante, no se ha logrado caracterizar a la LLA de linaje T de forma tan precisa y con un pronóstico claro asociado como ocurre en las de linaje B. En términos generales, la LLA-T se puede clasificar con base en las rutas biológicas implicadas, o en la expresión desregulada de factores de transcripción, mientras que los subtipos definidos por fusiones génicas son escasos. La activación oncogénica de factores de transcripción como *TAL1*, *TAL2*, *TLX1*, *TLX3*, *LMO1* y *LMO2* definen subtipos que en ocasiones guardan cierta relación con el estadio de diferenciación de linaje T (Liu et al., 2017; Inaba y Pui, 2021). Las alteraciones en *TLX1* y *TLX3* suelen presentarse en leucemias T en estadio de

maduración cortical temprana, mientras que la activación de TAL1 es característica de etapas corticales tardías (Soulie et al., 2005; Mansour et al., 2014). La desregulación de factores de transcripción como LMO1, LMO2, y BCL11 suele estar presente en LLA de células precursoras T tempranas (ETP por sus siglas en inglés) (Coustan-Smith et al., 2009). ETP es un subtipo altamente agresivo caracterizado por la presencia de marcadores de células troncales, y posiblemente se origina en un progenitor multipotente en etapas muy tempranas de maduración (Coustan-Smith et al., 2009; Zhang et al., 2012). Otros genes frecuentemente alterados son *NOTCH1*, *CDKN2A/B* y *FBXW7* (Hu et al., 2017; Ma et al., 2018). Anteriormente, el linaje T era considerado por sí mismo un factor de pronóstico adverso. Los estudios recientes han permitido refinar la clasificación de la LLA-T e identificar subtipos con mejor pronóstico, como los dominados por TLX1 y TLX3, o emplear terapias dirigidas como inhibidores de cinasas que muestran eficacia en subtipos como ETP que tienen sobre-activación de la vía JAK-STAT (Mansour et al., 2014; Maude et al., 2015; Inaba et al., 2017).

Asociación entre subtipos genéticos con edad, pronóstico y etnicidad

Sin tratamiento, la LLA es una enfermedad letal cuya supervivencia no superaba el 10% hasta la década de los 60s (Gupta et al., 2015; Hunger y Mullighan, 2015). Durante las siguientes décadas la supervivencia mejoró sustancialmente, llegando a 60% en la década de los 70s y casi al 90% a finales del siglo XX (Smith et al., 2010; Trigg et al., 2008; Pieters et al., 2016). Este éxito se debió principalmente a la adaptación de los esquemas terapéuticos basados en características de riesgo. Estos complejos esquemas están diseñados para intensificar el tratamiento para los pacientes de mayor riesgo y con leucemias más agresivas, o disminuir su intensidad y evitando efectos tóxicos innecesarios en los pacientes con menor riesgo. La cantidad de variables que pueden ser incluidas para determinar el riesgo depende enteramente de la capacidad tecnológica de cada país o centro hospitalario. Los criterios mínimos para la clasificación de riesgo, implementados por el Instituto Nacional del Cáncer (NCI) de EE. UU., son la edad y

la cuenta leucocitaria (CL) en sangre periférica al diagnóstico (Larson, 2016). Los pacientes de riesgo habitual (RH), también llamado riesgo estándar) son aquellos de entre 1 a 9 años y una CL menor a $50 \times 10^9/L$. Los pacientes se consideran de riesgo alto (RA) cuando son menores de 1 o mayores a 10 años y/o una CL $>50 \times 10^9/L$ (Larson, 2016). Otras variables de pronóstico son la respuesta a la quimioterapia, que tiene como fin ajustar su intensidad y evitar recaídas. La respuesta a glucocorticoides es sumamente útil, pero actualmente el monitoreo de EMR, definido como la persistencia de células leucémicas después de la administración de uno o varios fármacos quimioterapéuticos, es el factor con mayor valor pronóstico en la actualidad (van Dongen et al., 2015; Kruse et al., 2020). La medición de EMR permite el establecimiento de más grupos de riesgo (riesgo bajo, riesgo habitual/intermedio, riesgo alto y riesgo muy alto) y modificar la intensidad del tratamiento en función de la respuesta del paciente (Steinherz et al., 1996; Schultz et al., 2007; Harvey et al., 2010). El RA asociado a los pacientes mayores de 10 años se atribuía parcialmente a una menor tolerancia al tratamiento (Pui et al., 2012; Pui & Evans, 2013). Como ocurre en otros tipos de cáncer, el pronóstico empeora a mayor edad, siendo los AYAs y los adultos mayores los que tienen una menor sobrevida (Pulte et al., 2009; Roberts, 2018). Ahora se conoce que esta caída en las tasas de sobrevida (llamado en inglés “survival cliff”) también se explica por la presencia de subtipos genéticos más agresivos en niños mayores (Roberts et al., 2014; Herold et al., 2017; Siegel et al., 2018). Las alteraciones genéticas con mejor pronóstico como las hiperdiploidías (>50 cromosomas) y la fusión *ETV6::RUNX1* están presentes en alrededor de la mitad de los casos de LLA pediátrica, y su patrón de incidencia coincide con el principal pico de incidencia de la enfermedad (2 a 5 años) (De Braekeleer et al., 2012; Mullighan, 2012). Otro marcador común, la fusión génica *TCF3::PBX1*, es considerado de buen pronóstico en la actualidad con intensificación del tratamiento (Kager et al., 2007). Las alteraciones genéticas conocidas de mal pronóstico como las hipodiploidías y *BCR::ABL1* son poco frecuentes en la infancia, pero aumentan gradualmente su frecuencia en adolescentes y adultos jóvenes (Talpaz et al., 2006; Roberts, 2018; Inaba y Mullighan, 2020). El patrón de incidencia de estos marcadores

comunes cambia conforme al riesgo o los grupos de edad: *BCR::ABL1* es más frecuente en niños mayores y AYAs, así como la presencia de subtipos recientemente descritos como *MEF2D-r*, *DUX4*, *ZNF384* o Ph-like (Gu et al., 2019; Inaba y Mullighan, 2020) (Figura 4).

Los estudios en grandes cohortes permitieron identificar que existe una asociación entre la edad, el riesgo, y la frecuencia de alteraciones genéticas con significancia pronóstica (Liu et al., 2016; Reshmi et al., 2017; Gu et al., 2019). Existe inclusive una relación entre etnicidad y la frecuencia de algunos subtipos genéticos como *ETV6::RUNX1*, que es menos frecuente en poblaciones con origen latino/nativo-americano. En México, *ETV6::RUNX1* se ha reportado entre el 7-11% de los casos, con frecuencias similares en población centroamericana (Jiménez-Morales et al., 2008; Pérez-Vera et al., 2005; Carranza et al., 2013; Bekker-Méndez et al., 2014; Juárez-Avendaño et al., 2020; Mata-Rocha et al., 2022; Lee et al., 2022).

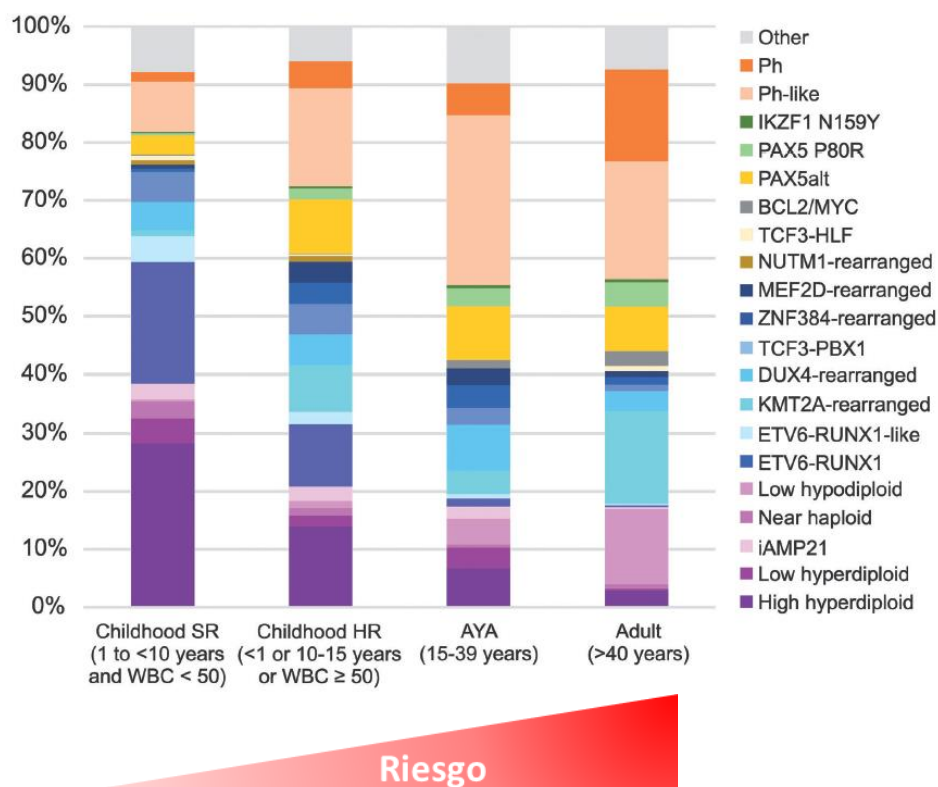


Figura 4. Distribución de subtipos genéticos en LLA-B respecto al riesgo y edad. SR: Standard risk; HR: High risk; WBC: White blood cell count; AYA: Adolescents and Young Adults. Tomada y modificada de Inaba y Mullighan, 2020.

Estas diferencias no parecen deberse a las tecnologías empleadas, ya que este patrón se observa también en niños con ascendencia latina (o “hispana”) en Estados Unidos (12% en latinos vs 20% en caucásicos) ([Kahn et al., 2018](#); [Lee et al., 2022](#)). Por otra parte, el subtipo Ph-like es más frecuente en pacientes con ancestría hispana/latina, especialmente alteraciones de *CRLF2* ([Harvey et al., 2010](#); [Roberts et al., 2014](#); [Reshmi et al., 2017](#); [Lee et al., 2022](#)).

Los estudios genómicos de gran escala como el pan-cáncer pediátrico nos ofrecen muchas pistas sobre la complejidad y gran diversidad de subtipos en LLA, pero la representatividad de diferentes poblaciones en ellos es aún muy escasa ([Ma et al., 2018](#); [Gröbner et al., 2018](#)). Esto es más relevante considerando las evidencias de que los patrones de incidencia y las características de riesgo en poblaciones menos representadas sugieren la presencia de casos con leucemias más agresivas.

Situación epidemiológica de la LLA pediátrica en México y el mundo

Varias regiones en el mundo han reportado un incremento en la incidencia de la LLA en décadas recientes, pero los países de alto ingreso continúan mostrando algunas de las tasas de incidencia más altas a nivel mundial ([Bonaventure et al., 2017](#)). Según reportes de la Agencia Internacional para la Investigación en Cáncer IARC, la incidencia de LLA en países europeos ronda entre los 30 a 40 casos por millón de niños aproximadamente ([Stiller, 2009](#)). En Asia, pocos países cuentan con registros epidemiológicos confiables, pero se estiman tasas de entre 32-41 casos/millón, principalmente en el sudeste asiático ([Demanelis et al., 2015](#)). La región con menor incidencia de LLA es la África subsahariana, con menos de 20 casos/millón, aunque la incidencia de LLA-T y algunos linfomas es mayor a la del resto del mundo ([Stiller, 2009](#); [Williams et al., 2014](#)). En el continente americano existe gran disparidad en las tasas de incidencia, con los registros de cáncer más completos en Canadá o EE. UU. pero inconsistentes en la mayoría de los países latinoamericanos. De acuerdo con el Programa Nacional de Registros de Cáncer, que incluye el 98% de la población de EE. UU., la incidencia de LLA es de 34

casos/millón, aunque con importantes diferencias por etnicidad o región (Siegel et al., 2017). Las tasas de incidencia en países latinoamericanos como Costa Rica, Ecuador, Colombia y algunas regiones de Brasil están entre las más altas a nivel mundial (52.2, 48, 40 y 56.5 casos/millón, respectivamente) (Quiroz et al., 2019). Aunque en México no existe un registro a nivel nacional de cáncer infantil, esfuerzos multiinstitucionales han reportado incidencias de entre 45 a 50 casos/millón, similares a las de los países latinoamericanos mencionados (Pérez-Saldivar et al., 2011; Steliarova-Foucher et al., 2017). Las características demográficas de la población estadounidense, en donde casi el 45% pertenece a algún grupo minoritario, permite hacer comparaciones entre grupos con diferente origen o etnicidad. La incidencia de LLA de la población hispana estadounidense es la mayor (y en la que se ha observado un incremento acelerado) en comparación con caucásicos y afroamericanos (42.9, 34.2, 18.7 casos/millón, respectivamente) (Siegel et al., 2017). La incidencia en estados del Oeste de EE. UU. (~40 casos/millón), en donde hay un mayor porcentaje de latinos son de origen mexicano, es mayor comparada con afroamericanos y caucásicos en igualdad de nivel socioeconómico (Siegel et al., 2017; US Census, 2021). El patrón de incidencia a lo largo de la vida también parece cambiar en algunos países o regiones en Latinoamérica, con tasas más altas en niños mayores y adolescentes. Acorde a la IARC, la incidencia de leucemias linfoides en niños entre 5-9 años las tiene Ecuador (44.5 casos/millón), México (44.3 casos/millón, solo Ciudad de México), Colombia (43 casos/millón) y Costa Rica (42.1 casos/millón) (Steliarova-Foucher et al., 2017). Este patrón es similar en algunos estados con una alta proporción de hispanos/latinos en Estados Unidos como California (39.5 casos/millón), Nuevo México (39 casos/millón) y Texas (36.3 casos/millón). La tendencia se repite en el grupo de 10-14 años en Ecuador (34.6), México (31.8) y Colombia (28.9). Como comparación, la incidencia en el grupo de 10-14 años en los dos estados de Canadá con mayor incidencia es de alrededor de 20 casos/millón, muy por debajo de los ~30 casos/millón en latinoamericanos. El Programa de Vigilancia, Epidemiología y Resultados Finales (SEER por sus siglas en inglés) del Instituto Nacional del Cáncer en EE. UU. muestra que la incidencia en

hispanos/latinos se incrementa alrededor de los 8 años, hasta ser casi el doble en la adolescencia comparado con caucásicos (Figura 5) (Quiroz et al., 2019).

Estas disparidades entre diferentes poblaciones y países también están reflejadas en mayor mortalidad. Existen incluso diferencias entre hispanos con distinto país de origen, ya que las tasas de mortalidad de los hispanos/latinos “continentales” (de origen mexicano, centroamericano o sudamericano) son mayores que la de los hispanos del caribe (de origen cubano o puertorriqueño) (5.6 muertes/millón vs 2.7 muertes/millón, respectivamente) (Shoag et al., 2020).

El programa CONCORD, establecido para evaluar la sobrevida global por cáncer en el mundo, resalta las grandes disparidades a nivel mundial. La sobrevida global a 5 años de LLA en países europeos de alto ingreso es de más del 90% (Bonaventure et al., 2017). En América, la mejor sobrevida se alcanza en Estados Unidos y Canadá (>90%), Costa Rica (~80%) y Argentina (76%), mientras que las menores tasas de sobrevida están en Ecuador (50%) y México (~53%) (Allemani et al., 2018). Los registros locales en México han estimado una sobrevida global a 5 años apenas superior al 60% (Jiménez-Hernández et al., 2014; Muñoz-Aguirre et al., 2021).

A lo anterior hay que agregar que una proporción muy alta de niños mexicanos con LLA presentan características de alto riesgo al diagnóstico (45%-70% de los casos), mientras que en países de alto ingreso este porcentaje no suele superar el 25%-30% (Núñez-Enríquez et al., 2019; Muñoz-Aguirre et al., 2020). Como ha sido planteado previamente, algunos subtipos de mal pronóstico como *BCR::ABL1*, o los recientemente descubiertos Ph-like y “Otros B”, están asociados a características clínicas de alto riesgo, siendo más frecuentes en mayores de 10 años o en pacientes con elevada CL. También siguen demostrándose las diferencias en las frecuencias de subtipos genéticos entre diferentes poblaciones, como la baja frecuencia del marcador de buen pronóstico *ETV6::RUNX1* en latinos, o la mayor frecuencia de Ph-like/CRLF2 en casos con ascendencia hispana (Jiménez-Morales et al., 2008; Bekker-Méndez et al., 2014; Lee et al., 2022; Mata-Rocha et al., 2022).

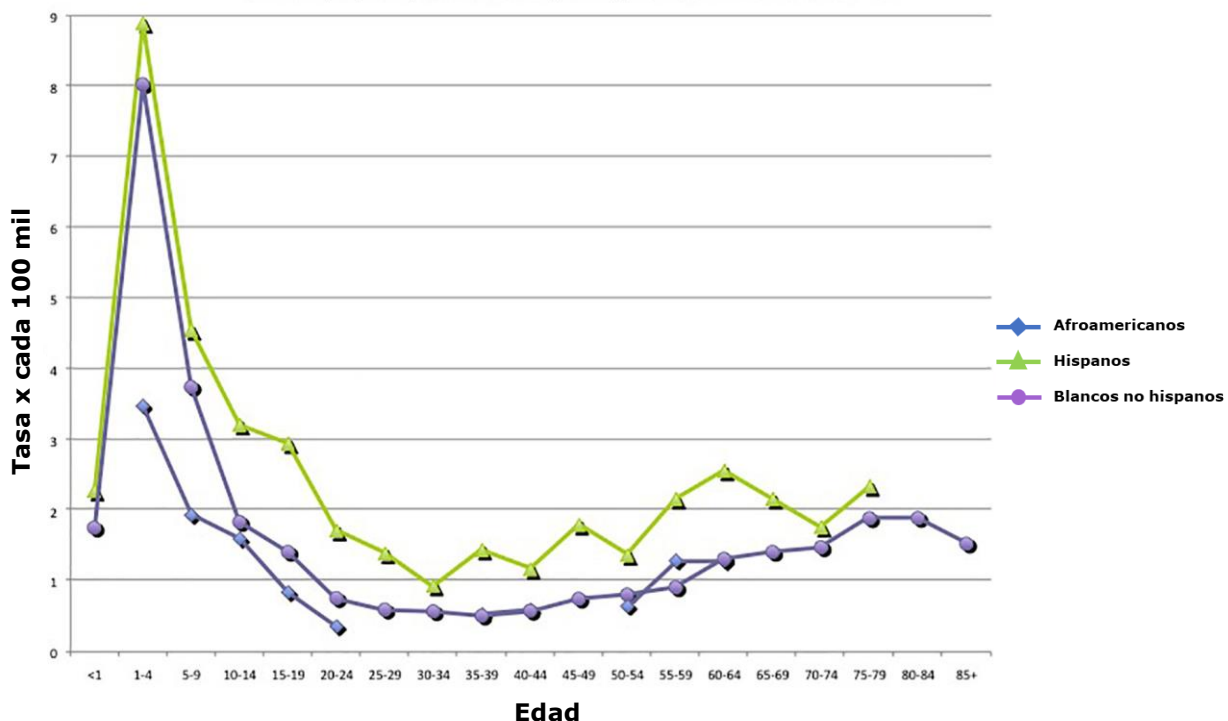


Figura 5. Tasas crudas de incidencia de LLA por edad respecto a la etnicidad. Datos del Programa de Vigilancia, Epidemiología y Resultados Finales (SEER) del Instituto Nacional del Cáncer en Estados Unidos. Tomada y modificada de Quiroz et al., 2019.

Los análisis genómicos nos han permitido conocer la biología de la enfermedad con un detalle sin precedente en el estudio del cáncer, pero la inclusividad de otras poblaciones no caucásicas en estos estudios es aún muy limitada, y es una barrera para aprovechar este conocimiento para el beneficio de ellas. El hecho de que México sea el cuarto país a nivel mundial con mayor pérdida de años de vida sana por LLA lo hace un problema de salud prioritario ([GBD 2017 Childhood Cancer Collaborators, 2019](#)). Es imprescindible realizar esfuerzos colaborativos para llevar a cabo estudios que reflejen apropiadamente la complejidad de la LLA en nuestro país y mejorar las tasas de supervivencia empleando herramientas con utilidad demostrada.

PLANTEAMIENTO DEL PROBLEMA

La incidencia de LLA pediátrica en México es de las más altas a nivel mundial, con alrededor de 50 casos por millón de niños cada año (Rivera-Luna et al., 2013; Fajardo-Gutiérrez et al., 2006; Steliarova-Foucher et al., 2017). Con la asignación de tratamientos con base en el riesgo, el monitoreo de la respuesta a la quimioterapia, y la identificación de nuevos subtipos genéticos con significancia pronóstica se han alcanzado tasas de sobrevida de más del 90% en países de alto ingreso. Sin embargo, la sobrevida en México no ha mejorado en décadas, y es menor incluso comparada con otros países latinoamericanos con infraestructura y nivel socioeconómico similares (Muñoz-Aguirre et al., 2021). Algunos factores que impiden mejorar las tasas de sobrevida están las muertes por el tratamiento y las recaídas. En relación con esto, en México entre 15%-22% de las recaídas ocurren dentro de los primeros 36 meses después del diagnóstico, denominadas recaídas tempranas (5%-9% en países de alto ingreso) (Jiménez-Hernández et al., 2015; Oskarsson et al., 2016; Núñez-Enríquez et al., 2019; Jeha et al., 2021). Solo durante el primer año de tratamiento, mueren entre 10%-14% de los casos (apenas un ~2% en Estados Unidos) (Jiménez-Hernández et al., 2015; Martín-Trejo et al., 2017; Núñez-Enríquez et al., 2019; Jeha et al., 2019), mientras que al menos la mitad presentan características clínicas de alto riesgo al diagnóstico (Jiménez-Hernández et al., 2015; Martín-Trejo et al., 2017; Núñez-Enríquez et al., 2019). Aunque estos problemas son indudablemente multifactoriales y también se observan en otros países de bajo y mediano ingreso, ahora sabemos que la frecuencia de subtipos genéticos también varía entre poblaciones. Algunos marcadores genéticos de buen pronóstico (en particular la fusión génica *ETV6::RUNX1*) son menos frecuentes en México, y el conocimiento sobre la presencia de nuevos subtipos de alto riesgo es aún escaso (Juárez-Avendaño et al., 2020; Rosales-Rodríguez et al., 2021; Juárez-Velázquez et al., 2022; Mata-Rocha et al., 2022). En nuestra línea de investigación hemos identificado perfiles de expresión génica asociados a recaída temprana en niños mexicanos con LLA. Para identificar con mayor precisión las alteraciones genéticas asociadas a mal pronóstico, es necesario el empleo de

herramientas de análisis masivo como la secuenciación, debido a la vasta heterogeneidad molecular que caracteriza a la LLA y a que algunas alteraciones genéticas recién descritas son difícilmente detectables con herramientas moleculares convencionales. Esto nos va a permitir conocer las alteraciones asociadas con mayor frecuencia con la progresión de la enfermedad, identificar los subtipos moleculares con mayor relevancia pronóstica, y poder incidir en una clasificación de riesgo más apropiada que requiera modificar la intensidad del tratamiento, evitar las recaídas y disminuir la mortalidad por LLA en nuestra población.

PREGUNTA DE INVESTIGACIÓN

¿Cuáles fusiones génicas y subtipos moleculares están asociados a recaídas tempranas y muertes en pacientes pediátricos mexicanos con leucemia linfoblástica aguda?

HIPÓTESIS

La presencia de alteraciones genéticas (fusiones génicas) no recurrentes está asociada a recaída temprana y muerte en pacientes pediátricos mexicanos con leucemia linfoblástica aguda

OBJETIVO GENERAL

Identificar alteraciones genéticas asociadas a recaída y muerte temprana en pacientes pediátricos con leucemia linfoblástica aguda.

Objetivos específicos

1. Identificar fusiones génicas (transcritos quiméricos) en niños mexicanos con LLA mediante secuenciación de transcriptoma (RNA-seq)
2. Identificar las rutas biológicas asociadas a las fusiones génicas presentes
3. Definir subtipos moleculares en LLA con base en las fusiones génicas conductoras identificadas
4. Conocer si existe asociación entre los subtipos moleculares identificados con variables clínicas (edad, cuenta leucocitaria)
5. Determinar si existen subtipos asociados a recaída temprana y/o muerte

METODOLOGÍA

Selección de pacientes

Se realizó un estudio de cohorte multicéntrico retrospectivo, como parte del Grupo Mexicano Interinstitucional para la identificación de las Causas de la Leucemia en Niños (GMIICLN), en el que participan el Instituto Nacional de Medicina Genómica y 8 hospitales públicos de la Ciudad de México:

1. Hospital de Pediatría del Centro Médico Nacional Siglo XXI. Instituto Mexicano del Seguro Social (IMSS).
2. Hospital Infantil de México "Federico Gómez". Institutos Nacionales de Salud.
3. Hospital Juárez de México. Hospitales Federales de Referencia.
4. Unidad Médica de Alta Especialidad (UMAE) Hospital General "Dr. Gaudencio González Garza", Centro Médico Nacional "La Raza". Instituto Mexicano del Seguro Social (IMSS).
5. Hospital Pediátrico Moctezuma. Hospital de Referencia de la Ciudad de México.
6. Hospital General de México "Dr. Eduardo Liceaga". Hospitales Federales de Referencia.
7. Hospital General Regional No. 1 "Dr. Carlos McGregor Sánchez Navarro". Instituto Mexicano del Seguro Social (IMSS).
8. Centro Médico Nacional "20 de noviembre". Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE)

Para este estudio se incluyeron pacientes en edad pediátrica (1-17.9 años) que fueron diagnosticados con LLA *de novo* entre 2014-2017, atendidos en alguno de los hospitales participantes y sin tratamiento de quimioterapia previo. El diagnóstico de LLA fue confirmado por oncólogos y/o hematólogos basados en criterios morfológicos e inmunofenotípicos de las células leucémicas obtenidas por aspirado de médula ósea (MO). Se realizó seguimiento clínico de al menos 36 meses después del diagnóstico. Los datos clínicos y moleculares (sexo, edad, CL en sangre periférica, inmunofenotipo, porcentaje de blastos en MO, esquema de tratamiento, tiempo de seguimiento, estudio molecular de las fusiones génicas

comunes en LLA) se obtuvieron de los registros recabados por el GMIICLN (los fármacos de quimioterapia de cada esquema se describen en la tabla suplementaria 4). Se consideraron los criterios de clasificación de riesgo NCI: RH = pacientes entre 1 a 9.99 años y CL menor a $50 \times 10^9/L$, RA = mayores a 10 años y/o CL $>50 \times 10^9/L$. Se obtuvo consentimiento informado de los padres/tutores y asentimiento informado en niños y niñas mayores de 8 años. Se excluyeron pacientes que cursaran con algún síndrome (trisomía 21, síndrome de Turner) o con enfermedades concomitantes (LLA como segunda neoplasia, infecciones). Se eliminaron del estudio a los pacientes o cuyos padres/tutores decidieron abandonar el estudio, aquellos que no hayan completado el tratamiento, y pacientes cuyo material biológico de la muestra de MO (RNA total) fue insuficiente para la secuenciación de RNA.

Las principales variables clínicas evaluadas fueron recaída temprana (RT) y muerte durante el período de seguimiento. La RT se define como la presencia de blastos leucémicos ($\geq 5\%$ de blastos) en MO, o evidencia de blastos leucémicos en sitios extramedulares (sistema nervioso central, testículos) después de haber alcanzado la remisión al término de la inducción ([Hunger y Raetz, 2020](#); [Buchmann et al., 2022](#)) (Figura 6). Se consideró remisión a la presencia de $<5\%$ de blastos en MO post-inducción y ausencia de blastos leucémicos en líquido cefalorraquídeo. Se determinó falla a la inducción si el paciente no alcanzó la remisión clínica después de la primera fase de la quimioterapia.

Procesamiento de muestras para análisis de transcriptoma (RNA-seq)

Se realizó la extracción de RNA total a partir de aspirados de MO obtenidos al diagnóstico (previo a cualquier tratamiento) mediante el método de tiocianato de guanidina-fenol-cloroformo ([Chomczynski y Sacchi, 2006](#)) con el reactivo TRIzol (Invitrogen, ThermoFisher Scientific). La pureza y concentración del RNA extraído fueron determinadas con un espectrofotómetro NanoDrop One (ThermoFisher Scientific). La integridad del RNA fue evaluada con electroforesis capilar mediante chips de bioanализador Agilent 2100 (Agilent Technologies, Santa Clara CA, EE. UU.).

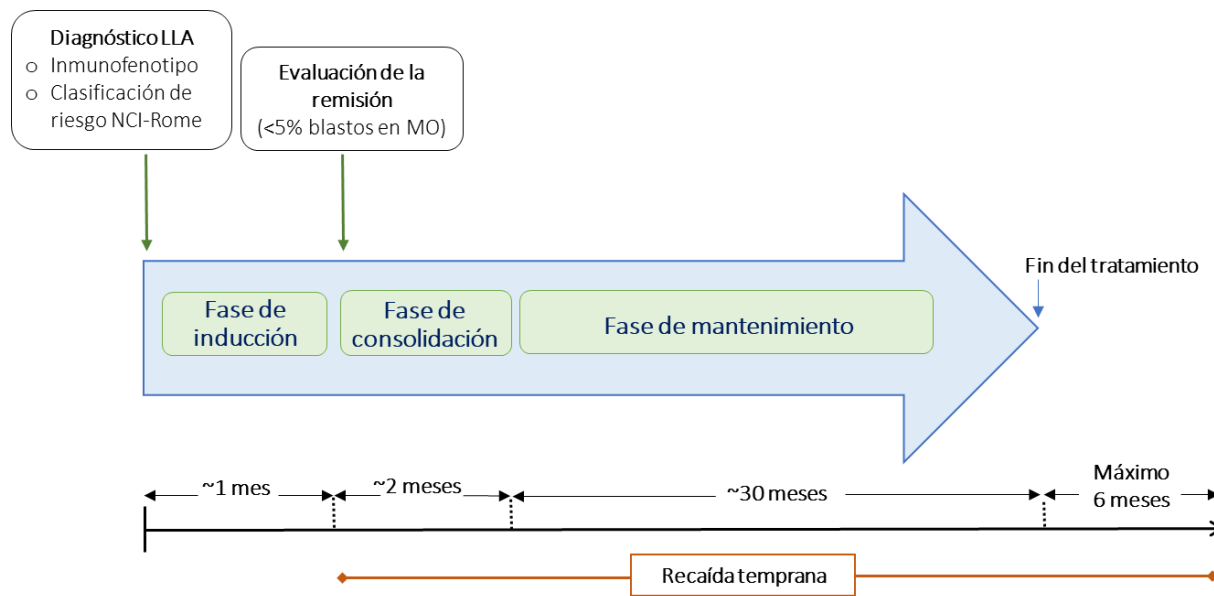


Figura 6. Fases del tratamiento de LLA. El tratamiento consta principalmente de 3 fases: inducción a la remisión, consolidación y mantenimiento. Las recaídas tempranas son aquellas que ocurren durante los primeros 36 meses a partir del diagnóstico (habiendo alcanzado la remisión post-inducción), o hasta 6 meses de finalizada la quimioterapia de primera línea.

Preparación de las bibliotecas de secuenciación

La secuenciación del transcriptoma fue realizada por la Unidad de Secuenciación del Instituto Nacional de Medicina Genómica. Las bibliotecas hebra-específicas (strand-specific) fueron elaboradas con el kit TruSeq Stranded Total RNA con depleción de RNA ribosomal (rRNA) (Ribo-Zero Gold) (Illumina, California, Estados Unidos) a partir de 1 microgramo de RNA total. Estas fueron secuenciadas en un secuenciador NextSeq 500 de Illumina, en lecturas de extremos pareados (paired-end) de 76 o 150 bases (PE 2x76 o PE 2x151). La profundidad de lecturas crudas obtenida fue de 47.3 millones en promedio por muestra (Rango: 7.2-92.9 millones).

Búsqueda de fusiones génicas con FusionCatcher

Para identificar la presencia de transcritos quiméricos/fusiones génicas se utilizó el programa FusionCatcher (versión 1.33) con parámetros default, con Python (versión 2.7.17) y BioPython (versión

1.59) como dependencias. La selección del programa se realizó con búsqueda en la literatura, en estudios en donde se haya comparado la sensibilidad y especificidad de distintas herramientas para identificar fusiones (Kumar et al., 2016; Tian et al., 2020). El programa fue instalado en un servidor de 6 núcleos con sistema operativo Ubuntu (18.04.3 LTS) y 126 GB de memoria RAM. FusionCatcher es un flujo de trabajo bioinformático (“pipeline”) diseñado para identificar fusiones/transcritos quiméricos en tejido somático con datos de secuenciación de RNA, preferentemente de bibliotecas de extremos pareados. FusionCatcher está implementado para procesar desde los datos crudos de secuenciación (rasurado de secuencias de baja calidad, adaptadores, secuencias no-humanas), y utiliza hasta 5 algoritmos de alineamiento distintos (BOWTIE, BLAT, STAR, BOWTIE2 y BMAP) (Nicorici et al., 2014) (Figura 7). La lista final de fusiones candidatas contiene los genes involucrados en cada evento, el número de lecturas de cada fusión, las coordenadas en el genoma del punto de unión entre ambos genes, y la secuencia predicha con el punto de ruptura. Los criterios mínimos respecto a número de lecturas fueron la presencia de al menos 3 lecturas que soportan la fusión (“spanning paired reads”) y al menos 2 lecturas que soportan el punto de fusión (“splitted reads” o “spanning unique reads”) (Figura 7). A partir de la lista de fusiones se realizó el primer filtrado para eliminar el mayor número de falsos positivos, con base en los términos/etiquetas provistas por el programa (Tabla suplementaria 1). Para eso, el programa reporta si la fusión es conocida, o si alguno de los genes está involucrado en cáncer en la literatura o bases de datos. También reporta si la fusión se ha identificado como falso positivo, fue identificada en individuos sanos, eventos “readthrough”, genes parálogos, pseudogenes, genes no codificantes, genes que codifican para RNAs estructurales como ribosomales, y RNAs pequeños (rRNA, tRNA, snoRNA). Transcritos altamente expresados, fusiones entre genes “promiscuos” (un gen se encuentra fusionado con múltiples genes más), genes con expresión constitutiva (ACTB, GAPDH), y fusiones en los que ninguno de los dos genes ha sido implicado en cáncer o no se conoce su relevancia biológica también fueron descartados (Tabla suplementaria 1).

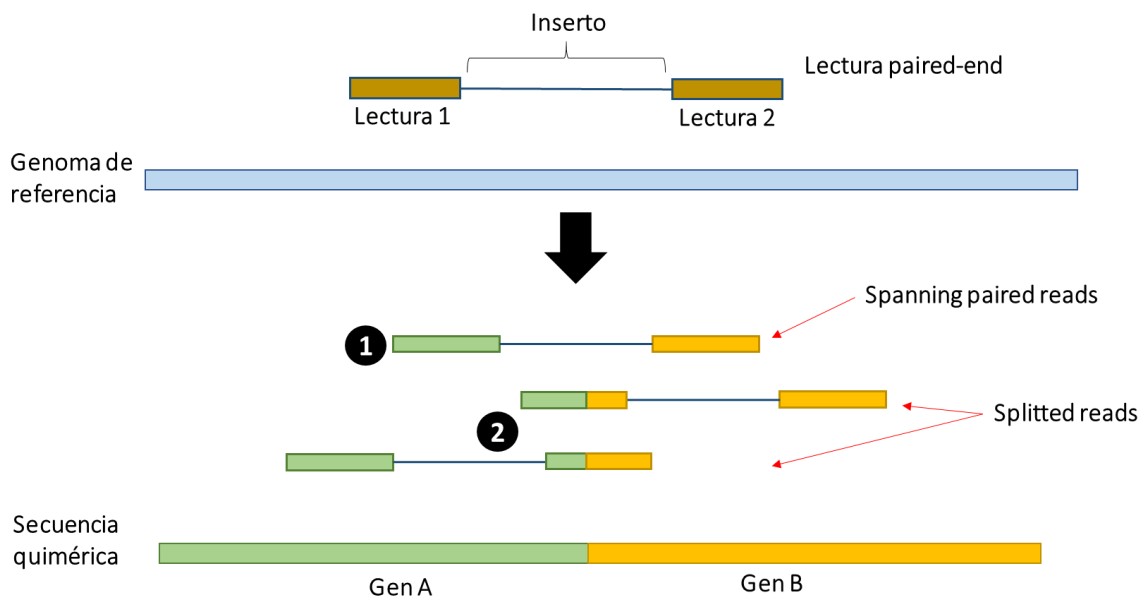


Figura 7. Conceptos básicos de RNA-seq para identificación de fusiones. Cada fragmento de secuenciación de extremos pares (biblioteca “paired-end”) está compuesto por dos lecturas (de longitud de 76 o 151 nucleótidos en este estudio), separadas por un fragmento llamado inserto, que no es secuenciado. Para identificar transcritos quiméricos las evidencias son: 1) La lectura 1 mapea al gen A y la lectura 2 mapea al gen B (“spanning paired reads”), los genes están físicamente distantes en el genoma de referencia, y 2) El punto de fusión de ambos genes se encuentra en una de las dos lecturas, produciendo una lectura quimérica (“splitted reads”)

Una vez eliminados los falsos positivos y filtrado por número mínimo de lecturas, el orden de los criterios para considerar fusiones fue el siguiente:

- 1) Fusiones/subtipos conocidos en LLA
- 2) La fusión no ha sido reportada previamente, pero uno de los genes involucrados en está frecuentemente alterado en LLA
- 3) La fusión no ha sido reportada en LLA, pero sí en la literatura/bases de datos en otros tipos de cáncer
- 4) La fusión mantiene el marco de lectura de ambos genes (“in-frame”)
- 5) La fusión interrumpe la secuencia codificante, o aumenta la expresión a partir del punto de ruptura predicho por el programa.

Todas las fusiones que hayan cumplido con los criterios mencionados fueron revisadas manualmente en el programa Integrative Genomics Viewer IGV versión 2.12 (Broad Institute) y en el alineador BLAT de

Ensembl (<http://www.ensembl.org/Multi/Tools/Blast>), usando GRCh38 como genoma de referencia (Robinson et al., 2011; Cunningham et al., 2022).

Para definir subtipos moleculares en LLA, una vez eliminados los falsos positivos, las fusiones se clasificaron como conductoras o concomitantes. Se consideraron fusiones conductoras (subtipo primario) si han sido descritas previamente y constituyen subtipos claramente definidos en LLA o leucemias de otros linajes (bifenotípicas), o si alguno de los dos genes involucrados es comúnmente mutado en LLA. En los casos en donde se identificaron fusiones concomitantes, la diferencia en el número de lecturas que soportan cada fusión fue indicativo de si se trataba de una alteración conductora o de un evento cooperador (Tabla suplementaria 2).

Búsqueda dirigida para identificación de fusiones del *locus IGH*

En los pacientes con LLA-B en quienes no se identificó ninguna fusión, se realizó una búsqueda dirigida para identificar fusiones que involucran al locus de la cadena pesada de las inmunoglobulinas (IGH), que ocasionalmente no son detectadas con RNA-seq. Los archivos crudos fueron mapeados al genoma de referencia GRCh38 con el programa STAR (Spliced Transcripts Alignment to a Reference) versión 2.7.2b, siguiendo el método de “2-pass” (Dobin et al., 2013). El archivo BAM de salida fue ordenado e indizado con Picard (Broad Institute) y samtools (HTSLib), respectivamente. Posteriormente, el archivo de alineamiento fue visualizado en el IGV para identificar lecturas rasuradas (“soft-clipped reads”), que consisten en fragmentos de lecturas que no pudieron ser alineados al genoma de referencia. Estas lecturas se buscaron en las regiones río arriba de *CRLF2* (~8 Kb río arriba del inicio de la transcripción), *DUX4* y *EPOR* (porción terminal del exón 8) para buscar el punto de ruptura de una probable fusión. El fragmento no alineado de la lectura fue ingresado a la herramienta en línea BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) buscando su alineamiento a cualquier región del *locus IGH* (o *IGK* e *IGK* como ha sido descrito en fusiones con *EPOR*). En caso de una probable fusión, se construyó el

cóntigo de la secuencia quimérica a partir del sitio de ruptura identificado buscando secuencias de entre 16-18 bases en los archivos fastq originales con el comando “grep” de UNIX.

Identificación de perfil de expresión Ph-like con datos de microarreglos

Debido al bajo número de pacientes por cada subtipo genético, nos apoyamos en una muestra más robusta para identificar casos con perfil de expresión Ph-like. Usamos datos de expresión génica con microarreglos (HTA 2.0 de Affymetrix) generados por nuestro grupo, provenientes de 72 pacientes con LLA. De estos pacientes, 31 forman parte de la población analizada con RNA-seq. El pre-procesamiento de datos crudos de intensidad de señal (archivos con extensión .CEL) se realizó en el programa Transcriptome Analysis Console (TAC 4.0 Affymetrix), usando un método de normalización por cuantiles con el algoritmo Robust Multi-array Average (RMA) (Irizarry et al., 2003). Se emplearon dos firmas de expresión para la identificación de Ph-like, validadas en grandes cohortes de microarreglos/secuenciación, compuestas por 8 (*JCHAIN*, *SPATS2L*, *CA6*, *NRXN3*, *ADGRF1*, *BMPR1B*, *MUC4*, *CRLF2*) o 15 genes (*BMPR1B*, *CA6*, *CRLF2*, *GPR110*, *JCHAIN*, *MUC4*, *NRXN3*, *SPATS2L*, *CHN2*, *IFITM1*, *PON2*, *S100Z*, *SEMA6A*, *SLC2A5*, *TP53INP1*) (Reshmi et al., 2017; Roberts et al., 2017). Ambas firmas de expresión son usadas para la clasificación de Ph-like en el Children’s Oncology Group de los Estados Unidos. Se realizó un análisis de agrupamiento jerárquico supervisado basándonos en 5 casos confirmados como *BCR::ABL1+* y 55 casos sin subtipo identificado (se descartaron *TCF3::PBX1*, *ETV6::RUNX1*, *TCF3::ZNF384*, *DUX4*, *MEF2D-r* para este análisis por ser usualmente excluyentes con Ph-like). También se emplearon los datos de microarreglos de expresión para evaluar la expresión de *CRLF2* (fenotipo Ph-like) y *DUX4* a lo largo de toda la cohorte, debido a que la activación de estos oncogenes es sugestiva de alteraciones genéticas (Lilljebjörn et al., 2016; Dou et al., 2017; Schmäh et al., 2017; Schinnerl et al., 2019). A partir de la matriz de expresión de los valores de intensidad normalizados, se seleccionaron las sondas correspondientes a los genes *CRLF2* (TCOX002305.hg.1 y TCOY000347.hg.1) y

DUX4 (TC18000271.hg.1 y TCUn_gl000228000014.hg.1) para identificar los casos con mayor expresión (cuartil superior) entre los 72 casos.

Validación de fusiones génicas

Las fusiones clasificadas como conductoras (drivers) y que definen subtipos en LLA fueron validadas con RT-PCR. Se realizó la síntesis de cDNA con transcripción reversa a partir de un microgramo de RNA total con el kit High Capacity cDNA Reverse Transcription (Applied Biosystems). Los oligonucleótidos para validar los transcritos quiméricos fueron diseñados en Primer-Blast (NCBI), con base en el punto de ruptura predicho por FusionCatcher y utilizando la secuencia del transcrito de Ensembl (www.ensembl.org/index.html) (Tabla suplementaria 3). *B2M* (β 2 microglobulina) y *ACTB* (actina B) fueron utilizados como genes controles para comprobar que hubiera material amplificable. Se realizó PCR de punto final en touchdown con la enzima Platinum Taq (Invitrogen, Life Technologies). Los amplicones fueron separados por electroforesis con un gel de agarosa al 2 o 2.5% y revelados en un transiluminador ChemiDoc Universal Hood III (Bio-Rad) con SYBR Gold (Invitrogen, ThermoFisher Scientific) como intercalante. Finalmente, los transcritos quiméricos no reportados previamente en la literatura y con material biológico disponible (RNA o cDNA) fueron validados con secuenciación Sanger.

Análisis estadístico

El análisis estadístico se realizó en R. Se utilizó ji-cuadrada o prueba exacta de Fisher para comparar entre variables clínicas categóricas, prueba de Kruskal-Wallis para CL y edad entre subtipos moleculares y prueba de Mann-Whitney-Wilcoxon para comparar grupos con sobre-expresión de *DUX4* y *CRLF2*. Se utilizó prueba de Kaplan-Meier para el cálculo de sobrevida global y sobrevida libre de evento.

RESULTADOS

Características clínicas de la población analizada con RNA-seq

Se seleccionó una subcohorta de 49 pacientes con LLA *de novo* para análisis de transcriptoma con RNA-seq (Tabla 1). Del total de pacientes, 49% fueron hombres y 51% fueron mujeres. Los pacientes fueron seguidos durante 31.2 meses (rango = 0.17 – 56 meses) en promedio. Debido a que la selección de pacientes se orientó hacia los que tuvieron mal pronóstico (que hayan presentado RT y/o muerte durante el seguimiento), se puede observar que la mayoría de los pacientes en esta población de estudio cuentan con características clínicas de RA bajo los criterios NCI. Veintinueve pacientes (59.2%) tenían más de 10 años al momento del diagnóstico (mediana = 10.9 años, rango = 1.5 - 17.3 años), y 45% presentaron una CL en sangre periférica mayor a $50 \times 10^9/L$ (mediana = $35 \times 10^9/L$, rango = 8 – $641 \times 10^9/L$). Diez pacientes (20.4%) presentaron hiperleucocitosis, definida como una CL superior a $100 \times 10^9/L$. En total, 71.4% se clasificaron como RA. Veintitrés (46.9%) pacientes fallecieron durante el período de seguimiento; de ellos, 17 (74% de las muertes) fallecieron durante los primeros 18 meses desde el diagnóstico. Dos pacientes fallecieron antes de completar la primera fase de inducción a la remisión. La supervivencia global de la población estudiada fue de 50% a 4 años de seguimiento (IC 95%: 37-67).

No se encontraron diferencias estadísticamente significativas respecto al sexo, edad, CL, inmunofenotipo y clasificación de riesgo entre los pacientes que presentaron RT en comparación con los que no la presentaron. Sin embargo, la mediana de la CL en sangre periférica fue incluso mayor en el grupo sin recaída reportada comparada con los que presentaron RT (41.6 vs $15.3 \times 10^9/L$, respectivamente).

Debido a que la recaída eleva sustancialmente el riesgo a morir, se observa que las muertes fueron más frecuentes en el grupo de pacientes con RT (69.5% vs 24%, $p=.004$) (Tabla 1).

La mediana de tiempo ocurrido a morir después de haber presentado una recaída fue de solo 2.8 meses.

La sobrevida global (SG) de los pacientes que presentaron RT fue de 29% (IC 95%: 15-57%) a 36 meses de seguimiento, muy baja en comparación a los casos que no recayeron (Figura 8).

Tabla 1. Características clínicas y moleculares de población de estudio

	Total n= 49	LLA-Sin recaída n= 26	LLA-Recaída temprana n= 23	<i>P</i> ***
	n (%)	n (%)	n (%)	
Sexo				
Masculino	24 (49)	11 (42.3)	13 (56.5)	.39
Femenino	25 (51)	15 (57.7)	10 (43.5)	
Edad (años)				
1 – 9.99	20 (40.8)	11 (42.3)	9 (39.1)	1
≥ 10	29 (59.2)	15 (57.7)	14 (60.9)	
CL. (x10⁹ / L)				
<10	11 (22.4)	4 (15.4)	7 (30.4)	.52
10-49.99	16 (32.6)	9 (34.6)	7 (30.4)	
50-99.99	12 (24.5)	8 (30.8)	4 (17.4)	
≥ 100	10 (20.4)	5 (19.2)	5 (21.8)	
Inmunofenotipo				
Linaje B	45 (93.9)	25 (96.1)	21 (91.3)	.59
Linaje T	3 (6.1)	1 (3.9)	2 (8.7)	
Fusiones génicas al dx*				
<i>ETV6-RUNX1</i>	5 (10.2)	0 (0)	5 (21.74)	
<i>TCF3-PBX1</i>	2 (4.1)	2 (7.7)	0 (0)	-
<i>BCR-ABL1</i>	3 (6.1)	2 (7.7)	1 (4.35)	
Sin fusiones recurrentes	39 (79.6)	22 (84.6)	17 (73.91)	
Esquema de quimio.				
St Jude Total XV	37 (75.5)	21 (80.7)	16 (69.6)	-
St Jude Total XIII y XV	3 (6.1)	1 (3.9)	2 (8.7)	
DFCI-00-01	1 (2)	0 (0)	1 (4.3)	
BFM-95	8 (16.3)	4 (15.4)	4 (17.4)	
Clasificación de riesgo**				
Riesgo alto (RA)	35 (71.4)	20 (76.9)	19 (82.6)	0.73
Riesgo habitual (RH)	14 (28.6)	6 (23.1)	4 (17.4)	
Muerte				
Sí	23 (46.9)	7 (26.9)	16 (69.5)	.004
No	26 (53.1)	19 (73.1)	7 (30.5)	

*Fusiones génicas recurrentes en LLA, identificadas como parte del estudio molecular realizado por el GMIICLN

**Criterios de clasificación de riesgo NCI (National Cancer Institute NCI, EE.UU.).

*** Valor de p solo en comparación LLA-RT vs LLA-sin recaída reportada (prueba de χ^2 o exacta de Fisher según el caso)

CL = Cuenta leucocitaria en sangre periférica

Adicionalmente, observamos que las recaídas que ocurren exclusivamente en MO (MO aislada), asociadas a peor pronóstico, se presentaron con mucha mayor frecuencia en pacientes que recayeron y fallecieron en comparación a los que recayeron, pero no murieron (81.25% vs 28.6%, respectivamente, $p=0.04$) (Tabla suplementaria 5).

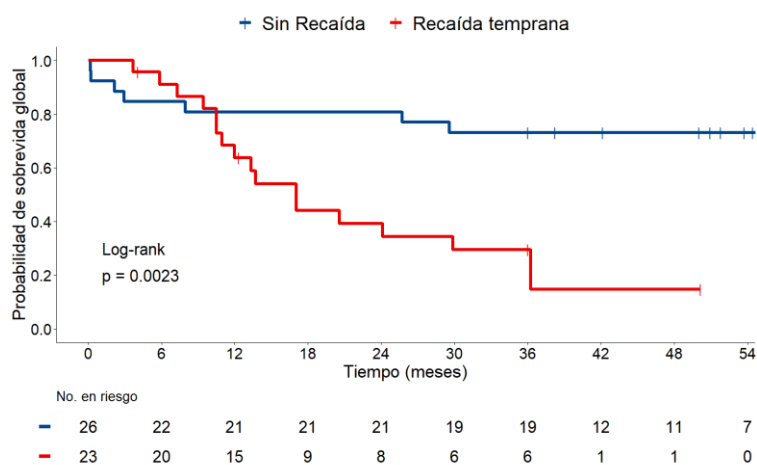


Figura 8. Sobrevida global con prueba de Kaplan-Meier de pacientes que presentaron recaída temprana

Pronóstico basado en características clínicas y clasificación de riesgo

Considerando las dos variables de la clasificación de riesgo NCI, edad y CL, no existieron diferencias en torno a la SG y la sobrevida libre de evento (SLE) a 36 meses de seguimiento entre ambos grupos de riesgo (RA: SG = 53.5%, IC 95% 39-73, SLE = 40%, IC 95%: 26-60; RH: SG = 53.8% IC 95% 32.6-89, SLE = 35.7% IC 95%: 17.7-72) (Figura 9).

Al analizar ambas variables de la clasificación de riesgo por separado, la SG de los pacientes mayores de 10 años fue inferior, aunque las diferencias no fueron significativas ($p=0.18$). La SG a 4 años de seguimiento en mayores de 10 años fue de 42% (IC 95%: 26-66%), en contraste con un 63% (IC 95%: 45-89%) en pacientes menores de 10 años a 36 meses de seguimiento (Figura 10A y 10B).

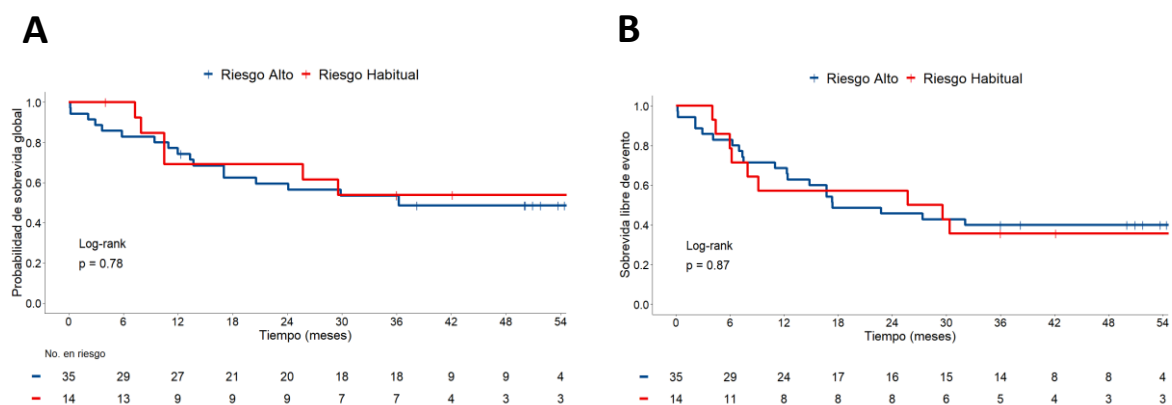


Figura 9. Probabilidad de supervida global (A) y supervida libre de evento (B) con prueba de Kaplan-Meier entre grupos clasificados de riesgo habitual (RH) y riesgo alto (RA) con criterios del NCI-

Por otra parte, al comparar dos grupos con base en la CL (punto de corte de $50 \times 10^9/L$ de la clasificación de riesgo) tampoco se encontraron diferencias significativas, aunque se observa una tendencia de menor SG y SLE en los pacientes que presentaron menos de $50 \times 10^9/L$ leucocitos en sangre periférica (Figura 10C y 10D). Interesantemente, al considerar un tercer grupo, separando los casos que presentaron hiperleucocitosis ($>100 \times 10^9/L$ leucocitos en sangre periférica), los pacientes que mostraron tendencia a una SG y SLE inferior fueron los de CL menor a $50 \times 10^9/L$ (SG de 50% y SLE de 33%) y los pacientes con hiperleucocitosis (SG de 50% y SLE de 30%) (Figura 10E y 10F). La hiperleucocitosis es por sí misma una variable pronóstica de muy alto riesgo, pero la cuenta inferior a $50 \times 10^9/L$ es una variable de buen pronóstico. En esta población de estudio, fueron los pacientes con CL entre 50 a $<100 \times 10^9/L$ los que tuvieron una mejor supervida (SG de 65% y SLE de 58%), lo que refleja los huecos existentes en la clasificación de riesgo cuando solo se consideran las variables clínicas presentes al diagnóstico.

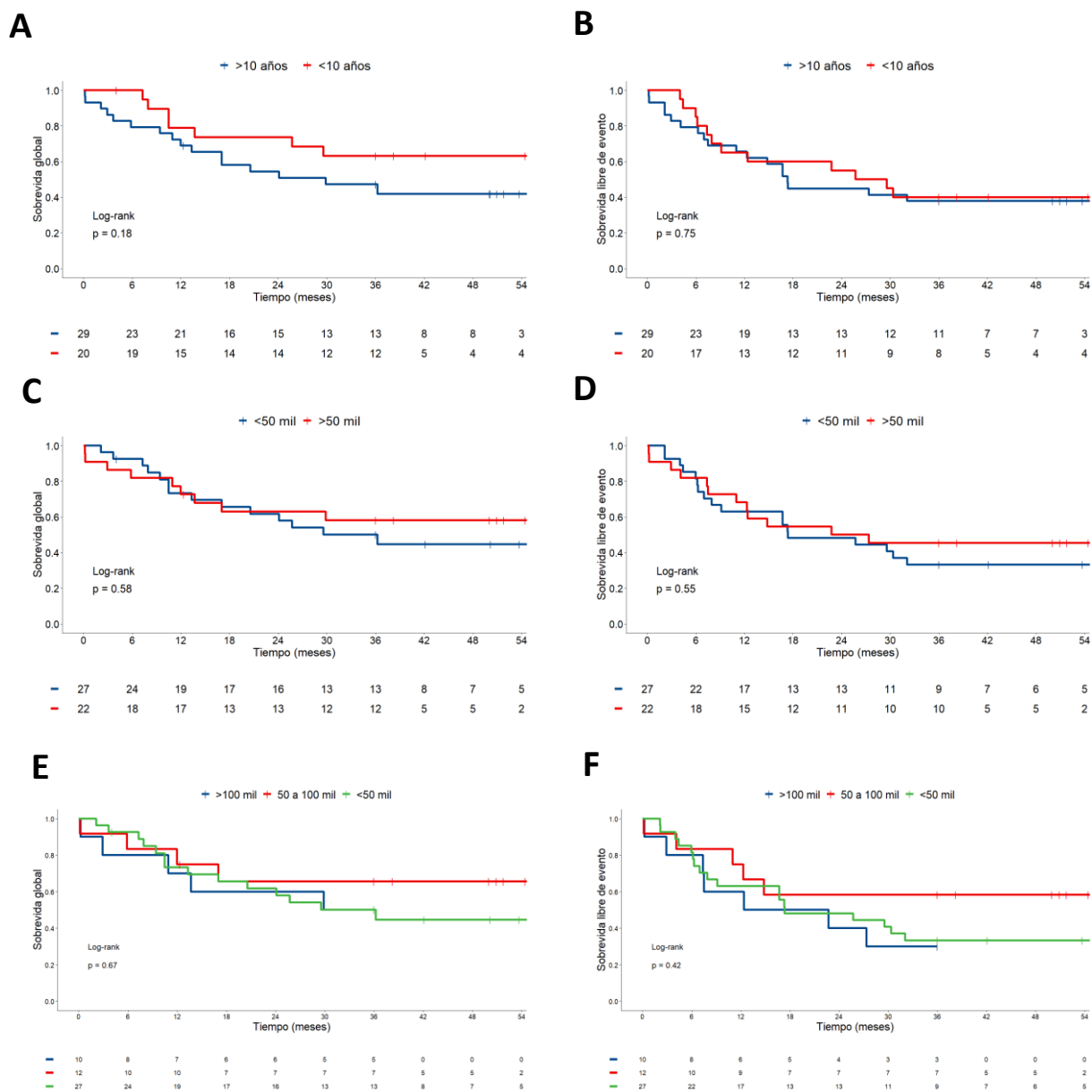


Figura 10. Probabilidad de sobrevivida global y sobrevivida libre de evento usando variables clínicas de riesgo. Valor de p de prueba de Kaplan-Meier para cálculo de sobrevivida global y sobrevivida libre de evento en función de la edad (A y B), cuenta leucocitaria en sangre periférica con valor de corte de 50 mil leucocitos/microlitro (C y D), y cuenta leucocitaria considerando un grupo más con hiperleucocitosis (E y F)

Fusiones génicas identificadas en LLA

De los 49 pacientes analizados, 31 presentaron al menos una fusión génica identificada con el programa FusionCatcher o con búsqueda dirigida (ver Materiales y Métodos) (Figura 11A y Tabla suplementaria 2).

Esto contrasta con los resultados de estudio molecular de rutina (con RT-PCR) realizado por el GMIICLN al diagnóstico del paciente, el cual mostró que 75% de los pacientes no contaban con ninguno de los marcadores más comunes, *ETV6::RUNX1*, *TCF3::PBX1*, *BCR::ABL1*. Se encontraron 30 fusiones génicas distintas, entre las cuales, más de la mitad involucran genes que codifican para factores de transcripción linfoides o cinasas (Figura 11B). En casi 20% de las fusiones están involucrados factores de transcripción de linajes no linfoides (Figura 11B), razón por la cual han sido reportadas en leucemias agudas bifenotípicas o de linaje mixto. Veintidós fusiones se encontraron una sola vez, mientras que las 8 restantes se encontraron en dos o más pacientes. Solo 10 pacientes (20.4%) tuvieron alguna de las fusiones génicas comunes en LLA-B. De estas, la más frecuente fue *BCR::ABL1*, presente en 5 pacientes. Cuatro pacientes tuvieron la isoforma *BCR::ABL1* menor (*BCR::ABL1m* o p190), la más común en LLA-B pediátrica, y un paciente presentó la isoforma *BCR::ABL1* mayor (*BCR::ABL1M* o p210), marcador patognomónico de la leucemia mieloide crónica en adultos pero muy poco frecuente en LLA. En nueve fusiones (18.3%) que no se habían reportado previamente en la literatura, uno de los dos genes involucrados en la fusión está frecuentemente alterado en LLA, como es el caso de *DUX4::ETV6*, *PAX5::ST18*, *PLPBP::SET* y *CDKN2A::TRDC*.

Las fusiones génicas constituyen eventos conductores del cáncer que suelen acompañarse de mutaciones puntuales, amplificaciones o deleciones como mutaciones conductoras, mas no de otras fusiones génicas. Consistente con esto, 38.7% de las fusiones fueron eventos únicos en el paciente contra un 22.4% en los que se encontró otra fusión, posiblemente tratándose de eventos cooperadoras para la leucemogénesis. Algunas de estas alteraciones cooperadoras como *IGH@::CCND3*, *IGH@::BCL2* y *WWOX::IGH@* no se habían identificado previamente en LLA pero sí en otros tipos de cáncer no hematológico o en leucemias de otro linaje (principalmente en LMA y LALM).

Aunque múltiples estudios han identificado alteraciones estructurales asociadas a LLA de linaje T, estas son poco frecuentes y no hay una clasificación genética clara con pronóstico asociado como en la LLA-B.

En nuestra población de estudio solo 3 pacientes fueron LLA-T, y solo en uno de ellos se identificaron fusiones (*RAF1::TMEM40* y *KDM4C::PLEKHO1* en el mismo paciente). Finalmente, en 17 pacientes no se identificaron fusiones génicas con ninguno de los dos métodos empleados (FusionCatcher o búsqueda dirigida).

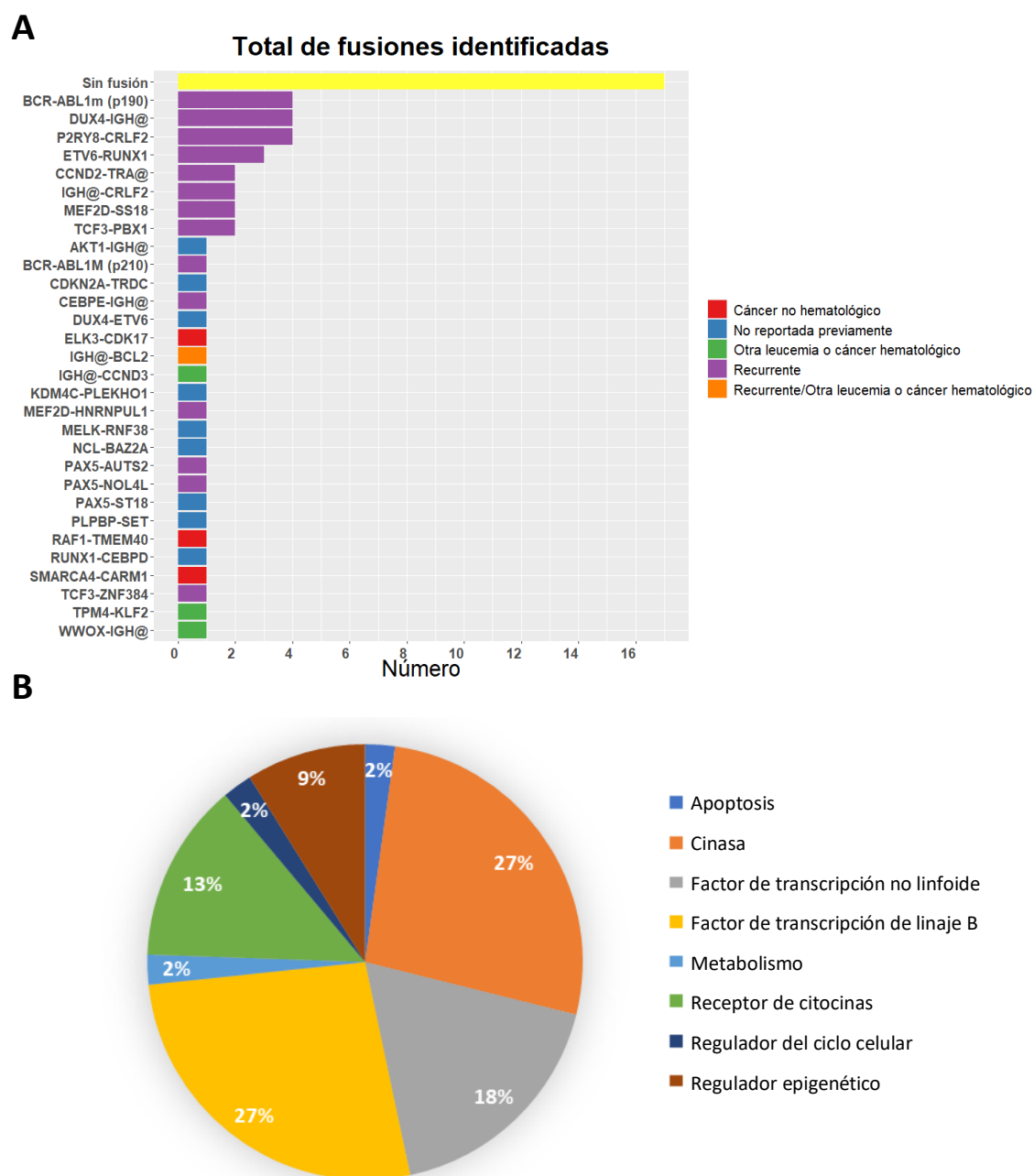


Figura 11. Fusiones identificadas en LLA pediátrica y funciones biológicas asociadas. A) Total de fusiones identificadas mediante secuenciación de transcriptoma RNA-seq. B) Papel biológico de los genes involucrados en las fusiones identificadas.

Conjuntamente, 10.9% de los casos de LLA-B fueron *BCR::ABL1+*. Es importante señalar que, de los 5 pacientes *BCR::ABL1+*, solo en dos de ellos se identificó la fusión en el estudio molecular de rutina realizado con RT-PCR. Los otros marcadores comunes en LLA-B, *ETV6::RUNX1* y *TCF3::PBX1*, se encontraron en 6.5% y 4.3% de los casos de LLA-B, respectivamente. En dos de los casos *ETV6::RUNX1+*, la fusión fue recíproca (*RUNX1::ETV6* genera un producto no funcional sin relevancia biológica). Al igual que *BCR::ABL1m*, estas fusiones se evalúan rutinariamente en el hospital de referencia de este estudio. Sin embargo, solo en 1 de 3 casos *ETV6::RUNX1+* fue identificada la fusión con RT-PCR. Las isoformas halladas de estos tres marcadores clásicos, *ETV6::RUNX1*, *TCF3::PBX1* y *BCR::ABL1* (menor y mayor) tuvieron los mismos puntos de ruptura canónicos reportados en la literatura.

En 4 pacientes se identificó la fusión *DUX4::IGH*, el cual conduce a la activación oncogénica de *DUX4* por su traslocación a los potenciadores del locus *IGH*. En un paciente se identificó la fusión *DUX4::ETV6*, reportada previamente una sola vez en la literatura. Dentro del subtipo genético *CRLF2-r*, considerado también Ph-like, la fusión *P2RY8::CRLF2* se encontró en 4 pacientes y en dos casos la fusión *IGH::CRLF2*. De los dos casos *IGH::CRLF2*, un paciente la presentó concomitante a *P2RY8::CRLF2*, mientras que en el otro caso solo pudo identificarse con búsqueda dirigida pero no con el algoritmo FusionCatcher.

Finalmente, 6 pacientes presentaron una fusión conductora considerada dentro del grupo “Otros B” (B-others). Tres pacientes tuvieron fusiones con el gen *MEF2D* (dos casos con *MEF2D::SS18* y un caso con *MEF2D::HNRNPUL1*), dos casos con fusiones de *PAX5* (*PAX5::NOL4L* y *PAX5::ST18*) y un caso con la fusión *TCF3::ZNF384*.

Identificación del perfil Ph-like con datos de microarreglos

Menos del 2% de los casos con sobre-expresión de *CRLF2* no se consideran Ph-like. Debido al limitado número de pacientes por subtipo de la población analizada con secuenciación, y a que no hay un método

como estándar de oro para identificar Ph-like, empleamos dos firmas de expresión de 8 y 15 genes validadas por el Children's Oncology Group (COG) (Reshmi et al., 2017; Roberts et al., 2018) en 72 casos analizados con una plataforma de microarreglos. De ellos, 31 fueron analizados tanto en microarreglos como con RNA-seq (Figura 12A).

Ambas firmas de expresión muestran la presencia de un grupo que distingue a los *BCR::ABL1+* (5 casos de 5 en la firma de 8 genes y 4 casos de 5 de la firma de 15 genes), y 8 casos sin esta fusión (Figura 12B y 12C). Estos 8 pacientes Ph-like representan el 11.1% de los casos de la cohorte de microarreglos.

Pudieron detectarse fusiones de *CRLF2* en 3 de estos casos Ph-like (dos *P2RY8::CRLF2+* y 1 *IGH::CRLF2+*). Se pudo confirmar que la sobre-expresión de *CRLF2* es un buen marcador para identificar casos con posibles alteraciones con este gen (Figura 12D). Además, las firmas génicas de 8 o 15 genes permitieron identificar casos Ph-like independientemente de las diferencias técnicas (las muestras fueron procesadas en varios lotes), las características de las muestras (p. ej. porcentaje de blastos), sin verse afectada la reproducibilidad. Este análisis complementario también dio la pauta para buscar la alteración *IGH::CRLF2* de forma dirigida en un paciente que presentaba la segunda mayor expresión de *CRLF2* de toda la cohorte de microarreglos, pero cuya fusión no pudo ser detectada con FusionCatcher (Figura 12D).

Identificación de outliers con expresión de *CRLF2* y *DUX4*

La naturaleza repetitiva del locus *IGH*, las regiones *PAR1/2* (que albergan a *CRLF2*), y la región de repetidos *D4Z4* en los cromosomas 4 y 10 que contienen a *DUX4*, todas en localización telomérica, son notoriamente difíciles de mapear. Esto impide que en ocasiones algunos algoritmos no detecten estas alteraciones, con puntos de ruptura únicos en cada paciente. Como se mencionó previamente, un caso con la fusión *IGH-CRLF2* estaba dentro de los 7 con mayor expresión de *CRLF2* (dentro del percentil 90) de la cohorte de microarreglos, pero solo se pudo identificar con búsqueda dirigida (Figura 12D).

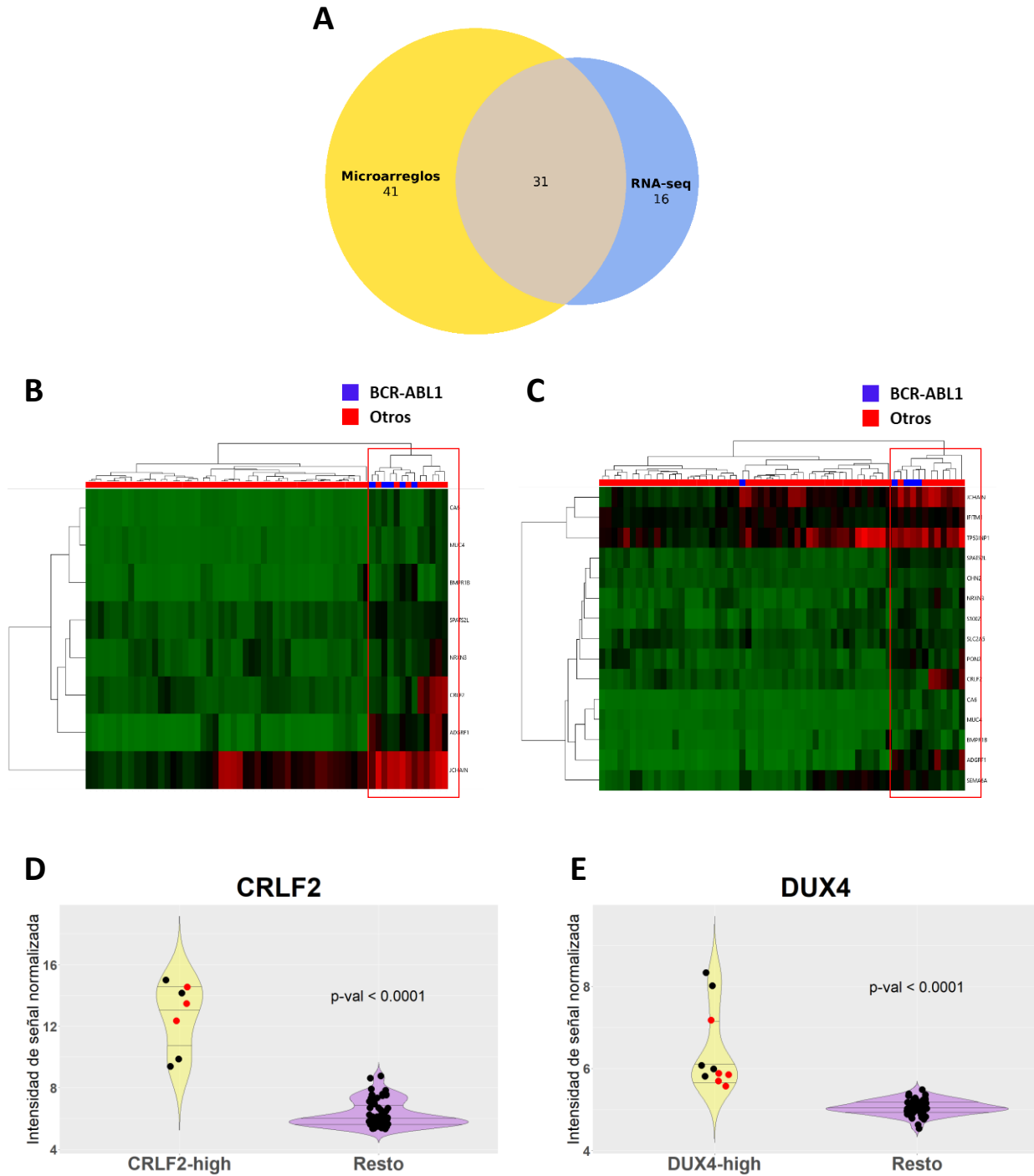


Figura 12. Identificación de subtipos moleculares con microarreglos de expresión. A) Número de muestras analizadas con microarreglos de expresión, RNA-seq, o ambas. B) y C) Perfil de expresión Ph-like con datos de microarreglos de expresión. Mapa de calor de los niveles de expresión de 8 genes (B) y 15 genes (C) correspondientes a firmas génicas para identificar Ph-like. Los casos *BCR-ABL1+* (azul) forman un grupo, en conjunto con 8 casos con perfil Ph-like. Entre los Ph-like, 3 pacientes presentaron alteraciones de *CRLF2* identificadas con RNA-seq (2 casos *P2RY8-CRLF2* y 1 caso *IGH-CRLF2*). Rojo = Sobre-expresión, verde = sub-expresión. D) y E) Figuras de violín de la expresión de los genes *CRLF2* y *DUX4* con datos de microarreglos de 72 muestras de LLA-B. Las muestras con el mayor nivel de expresión, denominadas *CRLF2-high* y *DUX4-high*, fueron comparadas contra el resto de los casos. Las líneas horizontales dentro del violín representan cuartiles y cada punto un paciente. Los puntos rojos representan casos analizados también con RNA-seq. Valor de *p* de prueba de Mann-Whitney-Wilcoxon

Esto confirma que la sobre-expresión de un oncogén, en este caso *CRLF2*, deriva de una lesión genética. Se realizó el mismo análisis para el gen *DUX4*, aunque las diferencias en nivel de expresión fueron menos evidentes que las que observamos en *CRLF2*. Se pudo asociar una mayor expresión de *DUX4* a una alteración genética subyacente identificada con RNA-seq: tres casos con mayor expresión de *DUX4* presentaron la fusión *IGH::DUX4* y un caso la fusión *DUX4::ETV6* (Figura 12E). Un paciente más presentaba uno de los mayores valores de expresión de *DUX4* (dentro del tercer cuartil, Figura 12E), pero no se identificó ninguna fusión con RNA-seq. Esta muestra tenía una baja profundidad (~17 millones de lecturas mapeadas) y la longitud de lecturas fue de 76 bases, factores pueden complicar la identificación de fusiones génicas. No obstante, dado que *DUX4* es un gen que no se expresa en tejido somático sano (De Iaco et al., 2017) y constituye un marcador de un subtipo plenamente identificado en LLA, se clasificó este paciente como parte del subtipo *DUX4*.

Clasificación por subtipos genéticos en LLA-B

Las fusiones génicas son eventos conductores del cáncer, especialmente en tumores pediátricos, y generalmente definen su comportamiento biológico. La activación de algunos oncogenes como *MEF2D* y *ZNF384* (en posición 5' o 3', respectivamente) mediante fusiones génicas define sus características biológicas al conservar sus dominios de activación transcripcional del extremo amino terminal, independientemente del gen con el que se fusiona. Esto hace posible clasificar a la LLA en subtipos con características particulares pese al gran número de alteraciones que podemos encontrar.

Diecisiete (34.7%) pacientes en este estudio presentaron solo una fusión, mientras que 14 (28.6%) presentaron al menos una fusión concomitante. En prácticamente todos los pacientes con alguna fusión concomitante es posible identificar el subtipo primario/alteración conductora por la diferencia en el número de lecturas entre ambas fusiones (Tabla suplementaria 2).

De esta forma, clasificando por subtipos moleculares recurrentes en LLA-B, los más comunes fueron *DUX4* (13%), *CRLF2* (10.87%) y *BCR::ABL1* (10.87%) (Figura 13). Dentro del subtipo *DUX4*, 4 pacientes presentaron alteraciones tipo *IGH::DUX4* (un caso con la fusión concomitante *ELK3::CDK17*), un caso la fusión *DUX4::ETV6*, y un paciente más se clasificó por sobre-expresión de *DUX4* determinada con microarreglos (Figura 12E).

Cuatro casos presentaron la alteración recurrente *PYRY8::CRLF2*, y en 2 de ellos se identificó alguna fusión concomitante (*IGH::CRLF2* y *CDKN2A::TRDC* en cada caso). De hecho, el caso *PYRY8::CRLF2* con la fusión concomitante *CDKN2A::TRDC*, presentó ambas alteraciones también al momento de la recaída (confirmado con RNA-seq y RT-PCR). El otro paciente *CRLF2-r* presentó la alteración *CRLF2::IGH* identificada mediante búsqueda dirigida de lecturas quiméricas.

La fusión recurrente *ETV6::RUNX1* estuvo presente en el 6.5% de los casos de LLA-B. Aunque esto puede deberse a las características de la población de estudio (con variables de riesgo alto), la baja frecuencia es consistente con lo reportado en otros estudios en México que demuestran la baja frecuencia de este marcador de bien pronóstico. En el mismo porcentaje (6.5% de LLA-B) se identificó el subtipo *MEF2D-r*, dos pacientes con la fusión *MEF2D::SS18* y uno con la fusión *MEF2D::HNRNPUL1*. Del paciente con la fusión *MEF2D::HNRNPUL1* también pudo obtenerse aspirado de MO a la recaída y la presencia de esta alteración también pudo confirmarse y asociarse a un evento de recaída.

El marcador *TCF3::PBX1*, también considerado recurrente, estuvo presente en un 4.3%, con la misma frecuencia del subtipo con rearrreglos de *PAX5* (un caso con la fusión *PAX5::ST18* y otro con la fusión *PAX5::NOL4L*). Aunque el subtipo *PAX5-r* (o *PAX5alt*) es clasificado principalmente con perfiles de expresión génica, hay un enriquecimiento en alteraciones con este gen, principalmente fusiones génicas con *PAX5* y algunas mutaciones puntuales (excluyendo la mutación *PAX5 P80R*). La fusión *PAX5::NOL4L* es de las más frecuentes dentro de *PAX5alt*, mientras que el gen *ST18* no había sido reportado

previamente como gen 3' unido con *PAX5*. El paciente con la alteración *PAX5::ST18* presentó otra fusión concomitante, *IGH@::BCL2*, reportada en LLA y en varios tipos de linfoma.

Subtipos moleculares en LLA-B

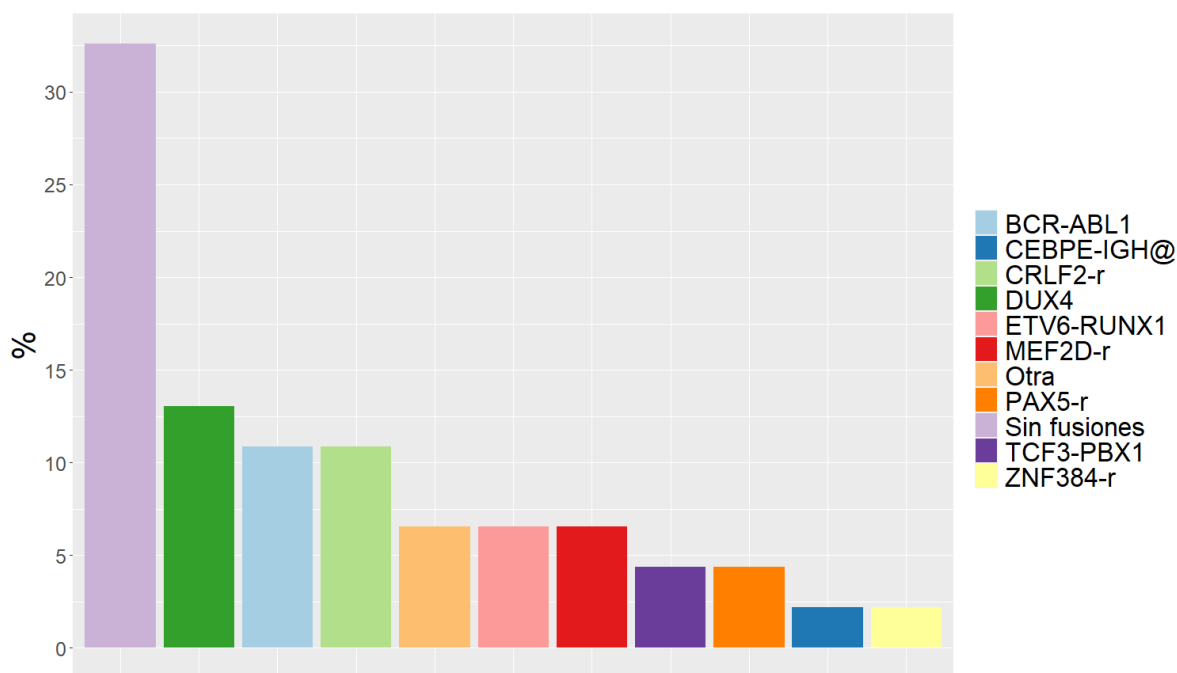


Figura 13. Porcentaje de subtipos genéticos identificados en LLA de linaje B.

Las fusiones que solo se encontraron en un solo paciente fueron *TCF3::ZNF384* y *CEBPE::IGH@*. De esta última, no es claro hasta ahora si constituye un subtipo en sí mismo, aunque frecuentemente su presencia es concomitante a hiperdiploidías o mutaciones de *ZEB2* (Gu et al., 2019; Brady et al., 2022).

En conjunto, 6.5% de los casos de LLA-B presentaron fusiones que no representan un subtipo reconocido, pero sí se han descrito en LLA, en otras neoplasias hematológicas, o en otros tipos de cáncer. En un paciente se identificó la fusión *PLPBP::SET*, quien además presentó otra fusión concomitante (*NCL::BAZ2A*). Ninguna de las dos fusiones ha sido previamente reportada, aunque existen reportes sobre fusiones con el gen *SET* con otros genes “partners” en LLA y en leucemias no diferenciadas/bifenotípicas. Un paciente presentó la fusión *TPM4::KLF2*, reportada en leucemia

promielocítica aguda, LMA y LLA-B, mientras que otro caso presentó la fusión *SMARCA4::CARM1*, que no ha sido previamente reportada en ninguna neoplasia hematológica pero sí en distintos tipos de adenocarcinomas (mama, ovario, estómago, pulmón) y en astrocitomas.

Características de las fusiones identificadas

MEF2D

Ambas fusiones del subtipo *MEF2D* halladas en este estudio son comunes en LLA-B, con los puntos de ruptura canónicos reportados en la literatura y conservando el marco de lectura. Las fusiones de *MEF2D* conservan los dominios de activación transcripcional de su extremo amino terminal, especialmente el dominio MADS, responsable de la unión a DNA. Las proteínas de fusión de *MEF2D* poseen mayor actividad transcripcional sobre sus genes blanco que en su estado silvestre, posiblemente porque la proteína del extremo carboxilo terminal estabiliza la unión a DNA de *MEF2D*.

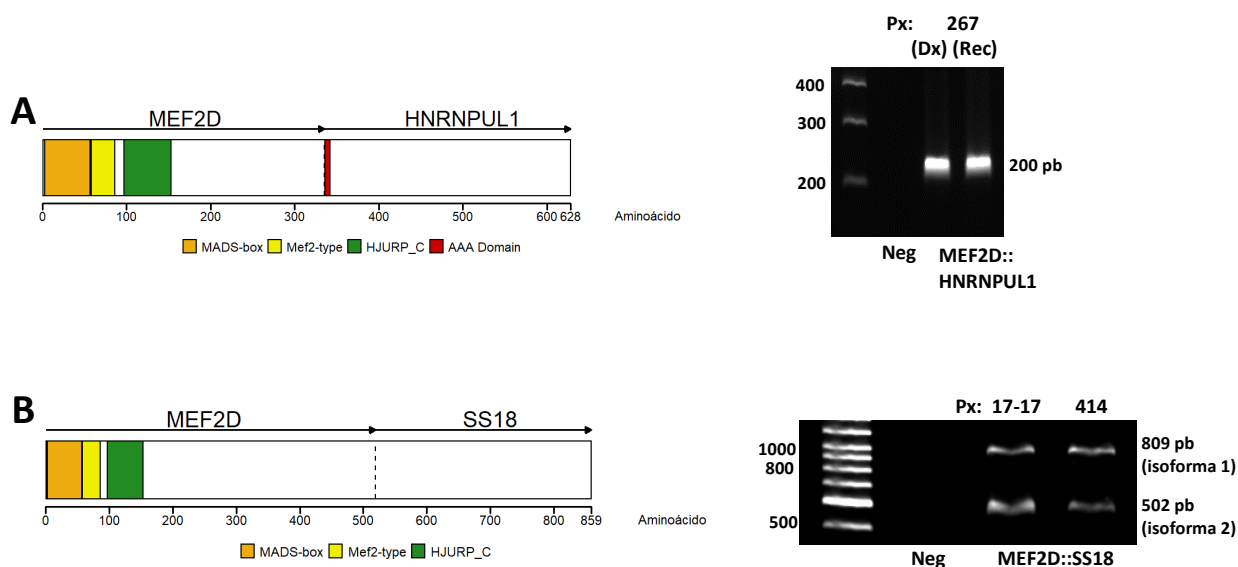


Figura 14. Fusiones del subtipo *MEF2D*. *MEF2D::HNRNPUL1* (A) se encontró en un solo paciente (ID: 267), el cual presentó recaída temprana. La fusión pudo confirmarse con secuenciación y RT-PCR en las muestras tomadas en ambos tiempos, al diagnóstico (carril 2) y a la recaída (carril 3). La fusión *MEF2D::SS18* (B) se identificó en dos pacientes (IDs 17-17 y 414). Se encontraron dos isoformas distintas para esta fusión, validadas con RT-PCR. Ambas isoformas de *MEF2D::SS18* difieren en el punto de ruptura de *MEF2D*, mientras que en *SS18* el punto de ruptura es el mismo.

La fusión *MEF2D::HNRNPUL1* genera un transcrito quimérico que une los primeros 9 exones de *MEF2D* y los exones 12-15 de *HNRNPUL1*. Esta fusión se encontró en un solo paciente, el cual presentó recaída y la misma fusión fue identificada al diagnóstico y a la recaída mediante RNA-seq y RT-PCR (Figura 14A). Dos pacientes presentaron la fusión *MEF2D::SS18*, con dos isoformas distintas que difieren en el punto de ruptura de *MEF2D* (hasta el exón 10 u 11) y el mismo punto de ruptura para *SS18* (a partir del exón 4). Ambas isoformas fueron validadas con RT-PCR (Figura 14B).

PAX5-r

Dos pacientes presentaron fusiones que involucran al gen *PAX5*, *PAX5::ST18* y *PAX5::NOL4L* (Figura 15A y 15B). Las fusiones de *PAX5* preservan los dominios de unión a DNA, pero pierden los dominios de transactivación situados en su extremo carboxilo terminal como suele ocurrir en este subtipo (Figura 15A). *ST18* no había sido previamente reportado como gen en posición 3'. Esta fusión une a los primeros 5 exones de *PAX5* con el exón 10 de *ST18*, conservando el marco de lectura y los dominios dedo de Zinc de unión a DNA de *ST18* (Figura 15C). *ST18* posee funciones de gen supresor tumoral en otros tipos de cáncer. La fusión *PAX5::NOL4L* es una alteración recurrente de este subtipo. El punto de unión entre ambos genes es el mismo reportado en otros estudios, entre el exón 6 de *PAX5* y el exón 7 de *NOL4L*.

Una tercera fusión, *PAX5::AUTS2*, fue identificada en un paciente *BCR::ABL1+*. Aunque *PAX5::AUTS2* es una alteración conocida, el punto de ruptura difiere de lo reportado en otras cohortes y ocurre en los extremos UTR de ambos genes (3' UTR de *PAX5* y 5' UTR de *AUTS2*), por lo que no hay pérdida de la expresión de *PAX5*.

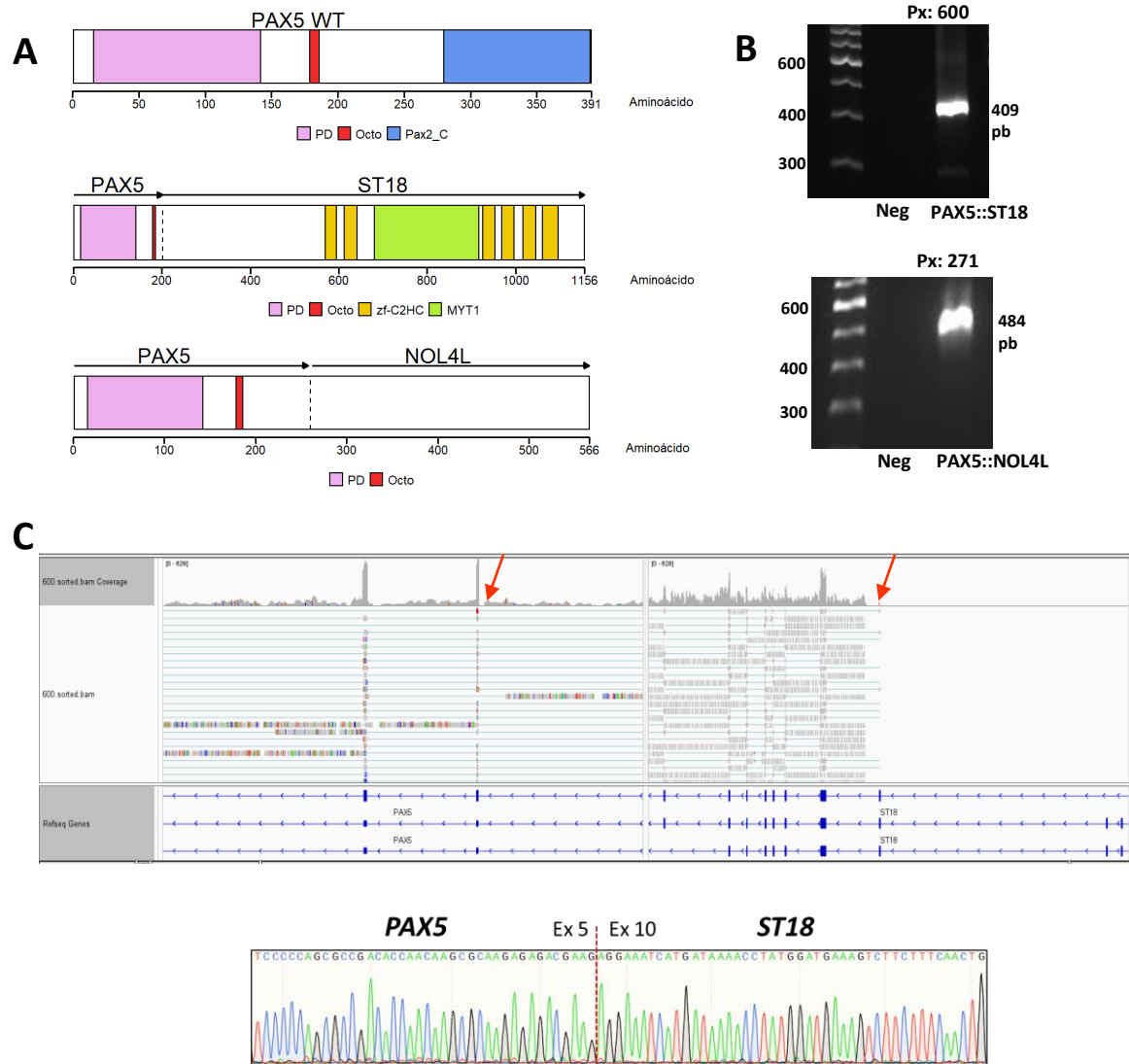


Figura 15. Fusiones del subtipo *PAX5-r*. A) Estructura de *PAX5* (wild-type, WT) y de las fusiones *PAX5::ST18* y *PAX5::NOL4L*. Las fusiones de *PAX5* pierden los dominios de regulación transcripcional de su extremo C-terminal. La fusión *PAX5::ST18* no ha sido previamente reportada, mientras que *PAX5::NOL4L* es recurrente en este subtipo. B) Productos de PCR de *PAX5::ST18* (ID: 600) y *PAX5::NOL4L* (ID: 271). C) Visualización de lecturas de la fusión *PAX5::ST18* en IGV y validación por secuenciación capilar. Las flechas rojas indican el punto de unión entre ambos genes. Se observa que la fusión lleva a la expresión de *ST18* a partir del exón 10. La secuenciación capilar confirma el punto de unión hasta el exón 5 de *PAX5*. PD: Paired domain; Octo: Octopeptide; Pax2_C: Paired-box protein 2 C-terminal; zf-C2H2: Zinc-finger C2H2-type; MYT1: Myelin transcription factor 1.

CRLF2-r

Los 4 casos con la fusión *P2RY8::CRLF2* presentaron la misma secuencia quimérica que une el primer exón no codificante de *P2RY8* con la secuencia entera de *CRLF2* (Figura 16A). Se identificaron 2 isoformas, difiriendo en la unión entre *P2RY8* con el primer o segundo exón de *CRLF2* (Figura 16B);

invariablemente, esto conduce a la sobre-expresión de *CRLF2* y la activación de genes con su dominio de fibronectina tipo III (Figura 16A).

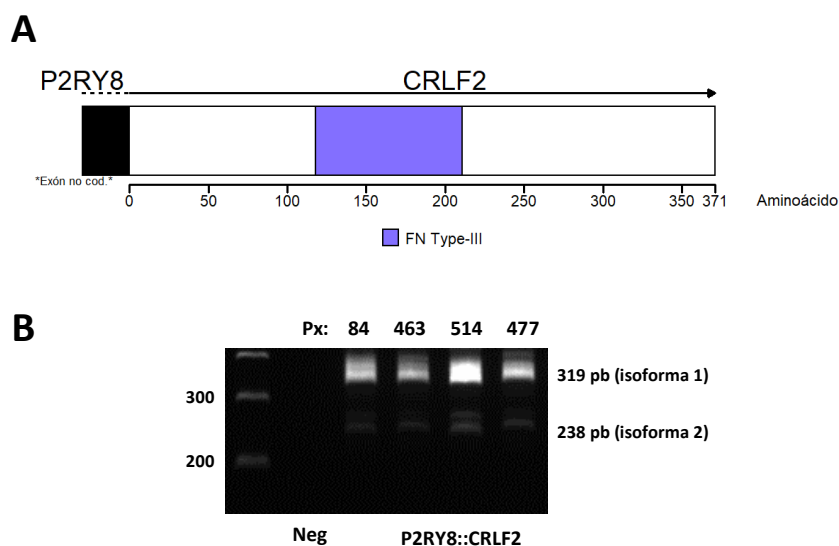
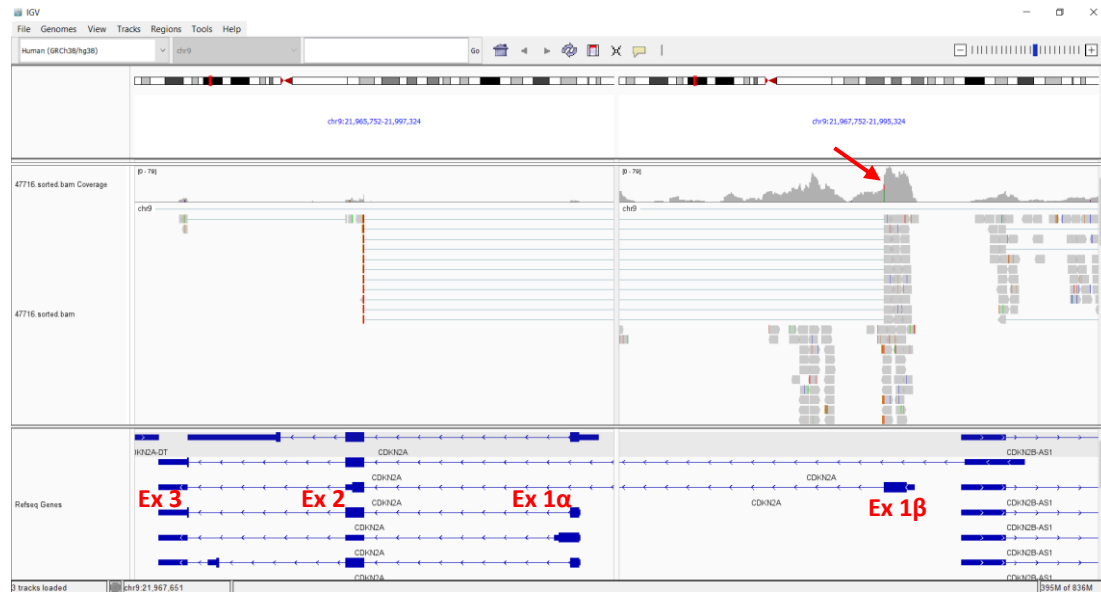


Figura 16. Fusión *P2RY8::CRLF2*. A) La fusión conduce a la sobre-expresión de *CRLF2* por la translocación de su secuencia entera al primer exón codificante de *P2RY8*, cuyos promotores son altamente activos en precursores linfoides. B) Productos de PCR de los 4 casos *P2RY8::CRLF2+* (Pacientes: 84, 463, 514 y 477). Se identificaron dos isoformas con FusionCatcher y RT-PCR, entre *P2RY8* y el exón 1 o 2 de *CRLF2*. FN Type-III: Fibronectin Type 3 Domain.

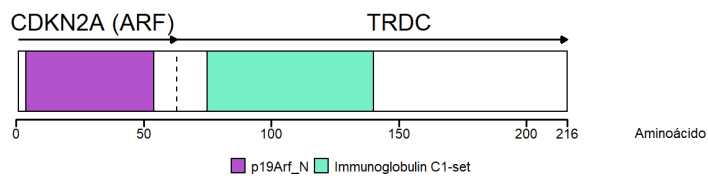
A pesar de considerarse una alteración conductora, en una minoría de casos *P2RY8::CRLF2* es subclonal y puede coexistir con otros subtipos. En nuestro estudio, dos pacientes *P2RY8::CRLF2+* presentaron alguna alteración concomitante, entre ellos uno con la fusión *IGH::CRLF2*, pero en una baja cantidad de lecturas (Tabla suplementaria 2). En otro paciente, la fusión concomitante *CDKN2A::TRDC* estuvo presente en un número similar de lecturas al de la fusión conductora *P2RY8::CRLF2*. El supresor tumoral *CDKN2A* es el gen mutado con mayor frecuencia en cáncer, aunque comúnmente por deleciones. De *CDKN2A* se producen dos transcritos con secuencia completamente distinta por la presencia de un primer exón alternativo, el exón 1 β que codifica para ARF (o p14ARF) y el exón 1 α para INK4a (o p16INK4a). La fusión predicha en este paciente une al exón 1 β de *ARF* con el gen del receptor constante delta de células T, *TRDC* (Figura 17B y 17C). Se puede observar que esta alteración abate por completo la expresión del

resto de la secuencia de *CDKN2A* (exones 2 y 3), incluyendo el exón 1 α de *INK4a* (Figura 17A). De este paciente se pudo obtener MO al momento de la recaída y ambas fusiones se identificaron con RNA-seq y RT-PCR. *P2RY8::CRLF2* y *CDKN2A::TRDC* se asociaron a un evento desfavorable (recaída temprana y muerte).

A



B



C

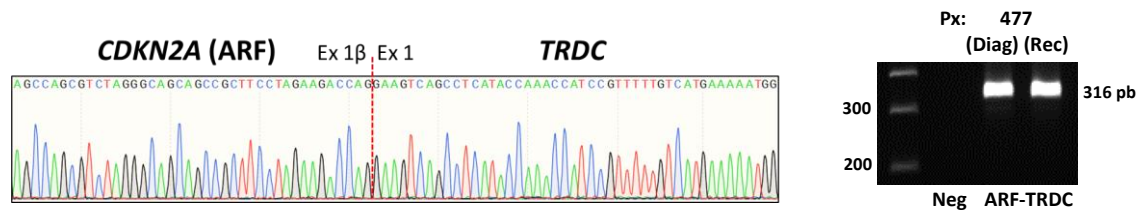


Figura 17. Fusión *CDKN2A::TRDC*. A) El panel derecho de IGV muestra el punto de ruptura (flecha roja) de *CDKN2A* en el exón 1 β (que codifica para el supresor tumoral *ARF*) que se une con el gen *TRDC*. La fusión interrumpe la expresión del resto de la secuencia de *CDKN2A*, incluyendo el exón 1 α que inicia la transcripción del inhibidor de cinasas *INK4a*. B) Modelo de la proteína de fusión *ARF::TRDC*. C) Secuenciación capilar y producto de PCR de la fusión. *ARF::TRDC* se encontró concomitante a la fusión conductora *P2RY8::CRLF2*. Ambas fusiones se encontraron tanto al diagnóstico como a la recaída del paciente en cuestión (la secuenciación capilar solo se realizó en la muestra al diagnóstico)

En un paciente, la fusión que une a *CRLF2* con el locus *IGH* se identificó mediante búsqueda dirigida de lecturas rasuradas (soft-clipped reads) (ver Materiales y Métodos, Figura 18). Como se indicó previamente, la expresión de *CRLF2* con los datos de microarreglos sugería la presencia de un rearrreglo génico que condujo a su expresión aberrante (Figura 12D). Estas lecturas se observaron a aproximadamente 6.3 kb río arriba del exón 1 de *CRLF2*. El mapeo del fragmento no alineado permitió detectar una fusión entre los RNAs potenciadores situados río arriba de *CRLF2* con una porción del gen *IGHJ6*, y este a su vez unido al gen de la cadena constante mu *IGHM* (Figura 18A y 18B). La presencia de un codón (GGT) insertado en la secuencia de *IGHJ6* pudo haber contribuido a no ser identificada por el algoritmo de fusiones, además de la presencia de 22 nucleótidos entre *CRLF2* e *IGHJ6* que no pudieron ser mapeados a ninguna región del genoma de referencia (Figura 18B y 18C). Este fenómeno es compatible con la adición de nucleótidos no templados característicos de actividad de la enzima desoxinucleotidil transferasa TdT (Clark, 1998), por recombinación aberrante observada en la generación de alteraciones que involucran a *IGH* (Iacobucci et al., 2016).

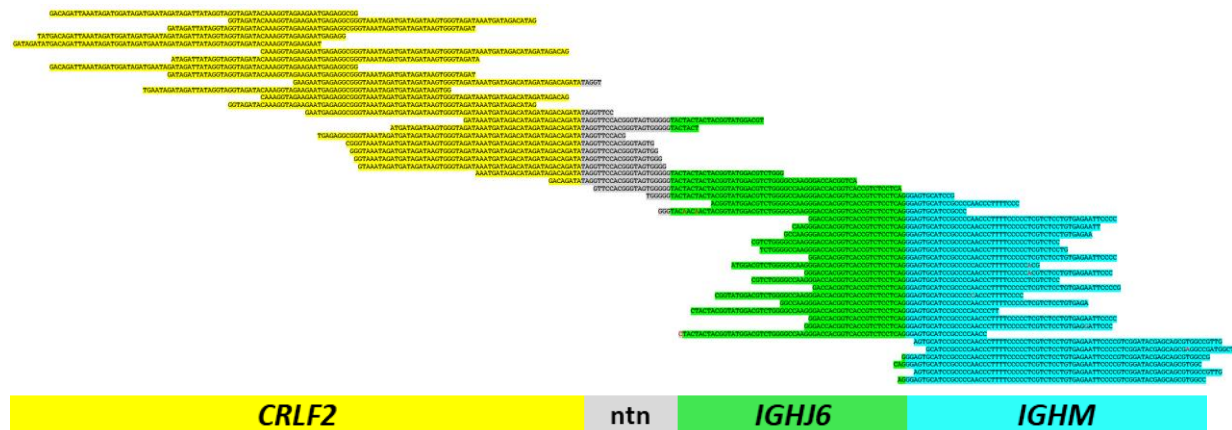
DUX4

De los 6 casos clasificados como *DUX4*, 4 presentaron la alteración más frecuente dentro de este subtipo que inserta al gen *DUX4* dentro de alguna región del locus de la cadena pesada de las inmunoglobulinas *IGH*. *DUX4* se localiza en una zona de repetidos en tándem llamados D4Z4 en la región subtelomérica del cromosoma 4, y en una región altamente homóloga en el cromosoma 10. Ambas están hipermetiladas en tejidos somáticos adultos y solo se expresa en testículos y etapas embrionarias tempranas. Cada repetido D4Z4 contiene un marco de lectura para un solo gen *DUX4*, cuya expresión es activada por su yuxtaposición (en cualquier dirección, 5' -> 3' o viceversa) en los promotores y potenciadores de *IGH* (Figura 19A).

A



B



C

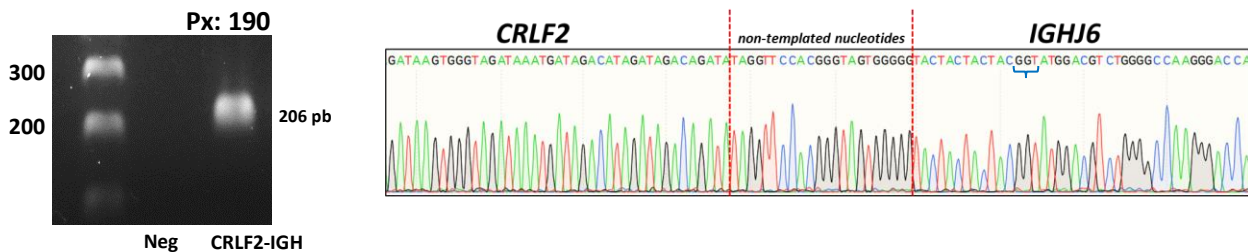


Figura 18. Identificación de fusión *CRLF2::IGH* con búsqueda dirigida. A) Visualización de lecturas en software IGV. El panel de la izquierda muestra la presencia de lecturas alineadas a aproximadamente 6.3 kb río arriba del sitio de inicio de la transcripción de *CRLF2* (cuadro amarillo). El panel de la derecha a mayor zoom muestra la presencia de lecturas rasuradas o “soft-clipped reads” (cuadros rojos) que fueron mapeadas al locus *IGH*. B) Construcción del cóntigo de la secuencia quimérica, compuesta por una porción del gen *IGHJ6* e *IGHM* en 3' y *CRLF2* en 5'. La presencia de nucleótidos sin templado sugiere recombinación aberrante y actividad de la enzima TdT. C) Producto de PCR (paciente #190) y secuencia del transcrito quimérico con secuenciación capilar. La llave azul indica el codón insertado en *IGHJ6*. ntn: Non-templated nucleotides

La gran variabilidad del locus *IGH* y la naturaleza repetitiva de D4Z4 (con alto contenido GC) dificulta la identificación de alteraciones *IGH::DUX4* incluso con secuenciación masiva.

En los 4 pacientes *IGH::DUX4* la inserción de *DUX4* ocurrió en sitios diferentes a lo largo del locus *IGH*. En un caso, *DUX4* se insertó entre las regiones *IGHJ* e *IGHD* (Figura 19B). En otro paciente, la inserción ocurrió entre los genes de la cadena constante de las inmunoglobulinas *IGHM* (también llamada cadena constante Mu) y aproximadamente 650 pares de bases río arriba del gen *IGHJ6* (Figura 19C). En otro caso la yuxtaposición de *DUX4* ocurrió en la porción más distal de la región *IGHV*, a ~1.2 kb de una secuencia potenciadora (Figura 19D). Finalmente, en otro paciente *DUX4* se insertó directamente con la secuencia codificante de *IGHM* (Figura 19E).

Como se mencionó previamente, solo en uno de los casos *DUX4* no se identificó la alteración genética subyacente, posiblemente debido a baja profundidad de lecturas. Esta limitación técnica se pudo resolver con los datos de expresión de microarreglos, que demostraron la sobre-expresión de *DUX4* que permitió clasificarlo dentro de este subtipo.

Un paciente presentó la fusión *DUX4::ETV6*, la cual solo se ha reportado una sola vez en la literatura. Esta fusión une al exón 1 de *DUX4* con el exón 2 de *ETV6* y conserva los marcos de lectura de ambos genes, incluyendo los dominios homeobox de *DUX4* encargados de la transcripción de sus genes blanco (Figura 20B y 20C). El mapeo de las lecturas en el programa IGV permite observar que la translocación de *DUX4* no interfiere en la expresión de los exones 2 en delante de *ETV6* (Figura 20A).

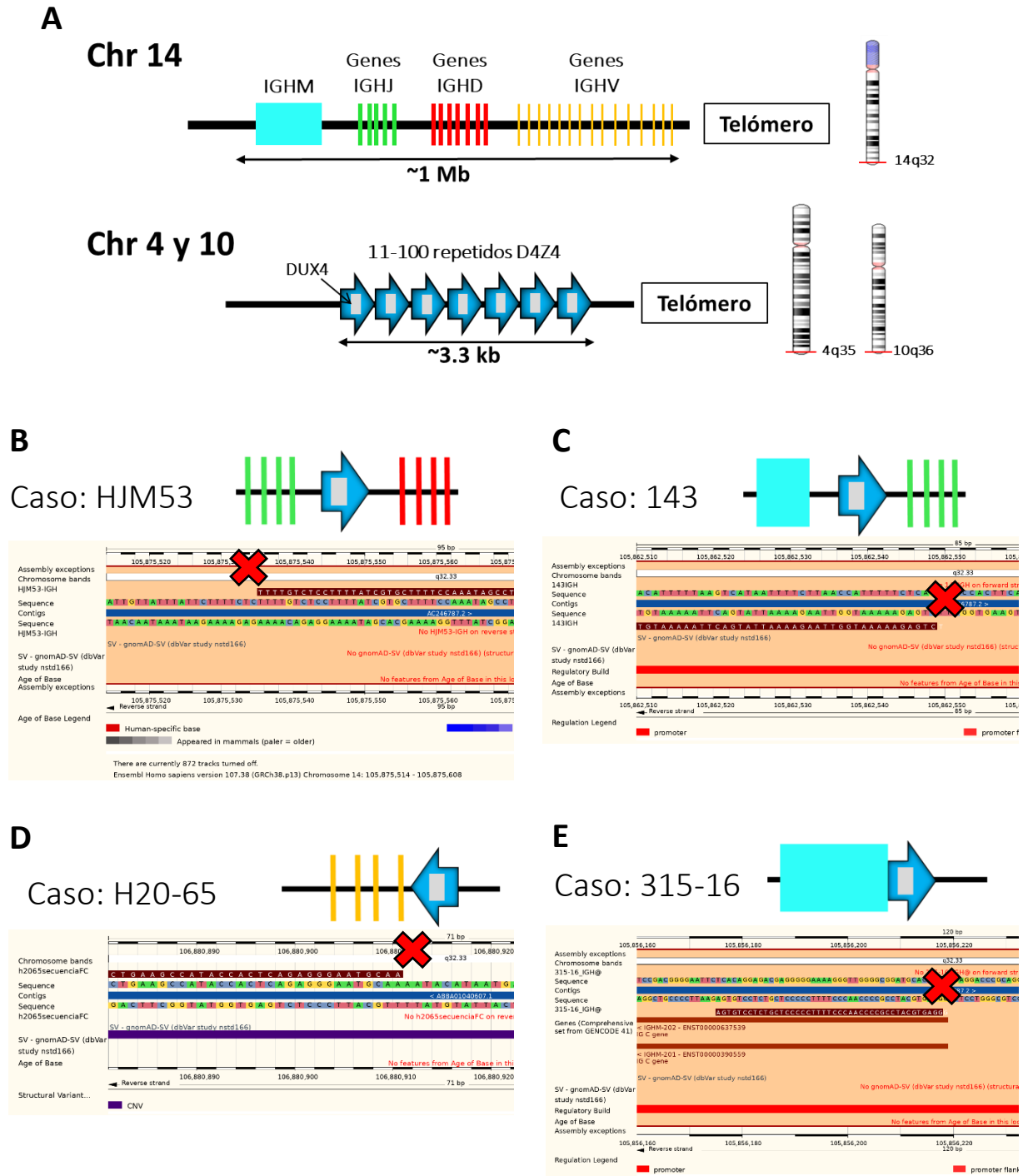


Figura 19. Fusiones *IGH::DUX4*. A) Estructura de *IGH* y *DUX4*. El locus *IGH* está localizado en la región subtelomérica del cromosoma 14 y está compuesto (de telómero a centrómero) por genes *IGHV* (variable), *IGHD* (diversity), *IGHJ* (joining) e *IGHM* (cadena constante mu). Los genes *DUX4* se localizan en una región de repetidos en tándem llamados D4Z4, situados en la región subtelomérica de los cromosomas 4 y 10. Los genes *DUX4* (un gen por cada repetido D4Z4) se pueden insertar en cualquier región a lo largo del locus *IGH*, lo cual conduce a su expresión. Los cuatro casos *IGH::DUX4* en este estudio presentaron arreglos únicos, con inserción de *DUX4* entre las regiones *IGHJ* e *IGHD* (B), entre *IGHM* e *IGHJ* (C), en la región más distal a *IGHV* (D), y adjacente al gen *IGHM* (E). La secuencia quimérica en la región *IGH* fue alineada con BLAT en la plataforma en línea de Ensembl y las cruces rojas denotan el punto de fusión con *DUX4*.

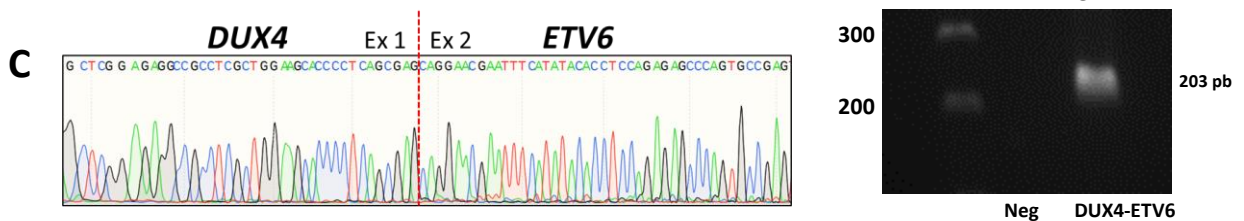
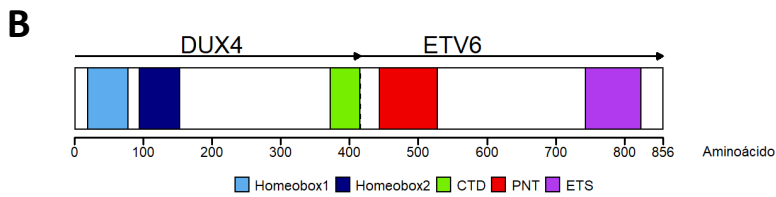
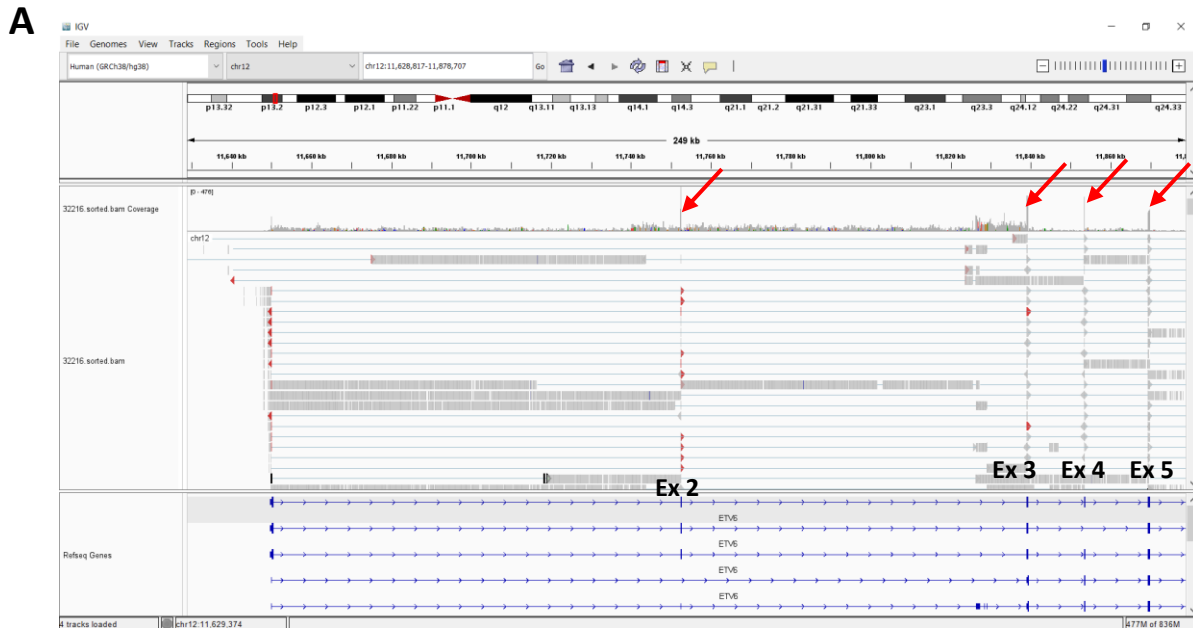


Figura 20. Fusión *DUX4::ETV6*. A) El alineamiento de las lecturas en IGV muestra que la expresión de *ETV6* a partir del exón 2 (flechas rojas) se mantiene a pesar de la inserción de *DUX4*. B) Modelo de los dominios conservados en la fusión, la cual conserva el marco de lectura de ambos genes, principalmente los dominios tipo homeobox de *DUX4* encargados de la transcripción de sus genes blanco. C) Producto de PCR y secuenciación capilar de *DUX4::ETV6*. CTD: C-terminal domain; PNT: Pointed domain of Tel-Yan; ETS: Erythroblast transformation specific domain.

ZNF384-r

Solo un caso presentó una alteración dentro de este subtipo, *TCF3::ZNF384* (Figura 21). La fusión ocurrió entre el exón 11 de *TCF3* y casi la secuencia entera de *ZNF384* desde su exón 2 (no codificante) (Figura 21B y 21C). En esta fusión, los dominios hélice-bucle-hélice de *TCF3* en su extremo carboxilo terminal (situados en 3') se pierden y se anula su actividad como factor de transcripción linfopoyético.

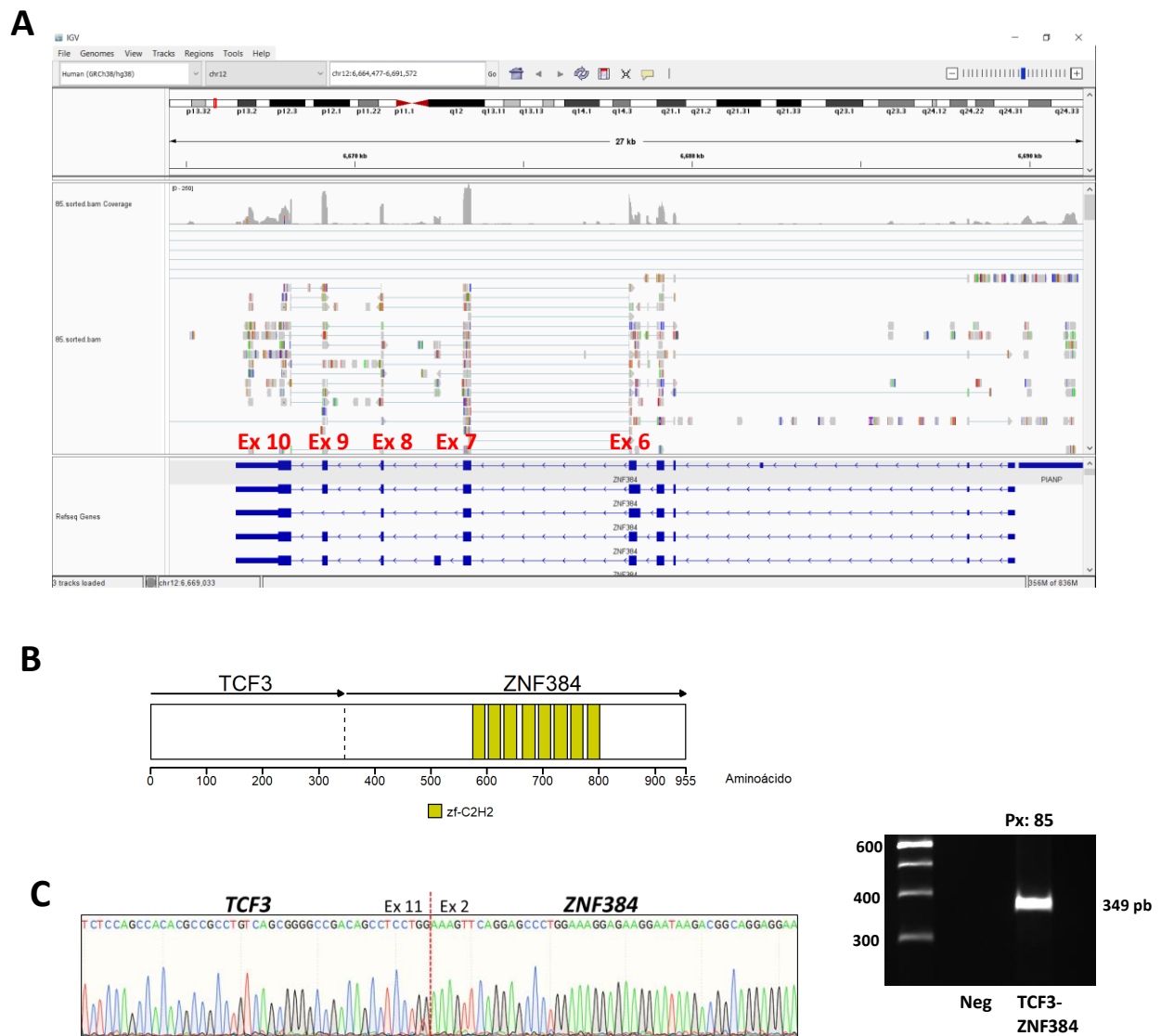


Figura 21. Fusión *TCF3::ZNF384*. A) Captura de IGV muestra la expresión de *ZNF384* desde el exón 2 (exón no codificante). B) Dominios de la proteína de fusión *TCF3::ZNF384*, en donde se conservan los dominios de activación transcripcional de *ZNF384*, localizados en los exones 6 a 10. C) Producto de PCR y de secuenciación capilar de *TCF3::ZNF384*. zf-C2H2: Zinc finger double domain

De *ZNF384* se conservan todos los dominios de tipo dedo de zinc, transcritos de los exones 6 y 10 (Figura 21A) y responsables de su actividad transcripcional y los que le otorgan las características biológicas de las fusiones tipo *ZNF384-r*. La fusión fue validada con secuenciación Sanger (Figura 21C) ya que la isoforma involucrada de *TCF3* no es la canónica o la más expresada de acuerdo con Ensembl.

Otras fusiones sin subtipo asignado en LLA

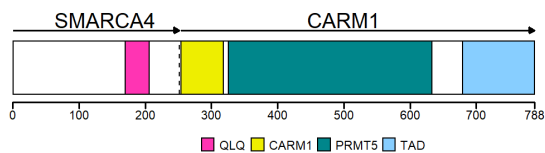
Cinco fusiones (*RAF1::TMEM40*, *SMARCA4::CARM1*, *TPM4::KLF2*, *PLPBP::SET*, *NCL::BAZ2A*) no pudieron ser asignadas a ningún subtipo conocido en LLA. En todas se conserva el marco de lectura y por lo tanto dominios funcionales, pero de ninguna se conoce su papel funcional. Algunas de estas fusiones han sido descritas en LLA, pero sin pertenecer a algún subtipo, o han sido implicadas en otras leucemias u otros tipos de cáncer. Solo en uno de 3 pacientes con LLA-T analizados se encontraron fusiones génicas, entre ellas *RAF1::TMEM40* (Figura 22D). Esta es la única fusión identificada en cáncer, en un solo caso reportado en LLA en bases de datos, en tumor de Wilms (tipo de cáncer renal más común en niños) y carcinoma de plexo coroideo (un tipo de tumor cerebral pediátrico poco común). Esta es una fusión intracromosómica, con una distancia de aproximadamente 160 kb entre ambos genes, que une los primeros 13 exones del proto-oncogén *RAF* con la secuencia codificante entera del gen de la proteína transmembrana 40 *TMEM40* (Figura 22D). La fusión predicha conserva los dominios de unión a RAS en su extremo amino terminal en 5' y una parte del dominio catalítico de *RAF*, además del único dominio funcional de *TMEM40* de su extremo carboxilo terminal, el cual tiene un posible rol oncogénico en algunos tipos de cáncer ([Ma et al., 2018](#)).

Por otra parte, alguno de los genes involucrados en las fusiones en los casos LLA-B sin subtipo asignado como *SMARCA4::CARM1*, *TPM4::KLF2* y *PLPBP::SET* (además de la fusión concomitante *NCL::BAZ2A*) tienen funciones en remodelación epigenética. La fusión *SMARCA4::CARM1* no ha sido previamente reportada en LLA, pero se han descrito mutaciones en *SMARCA4* recurrentes en LLA-T o involucrado en fusiones con *ZNF384* ([Dickerson et al., 2022](#)). El punto de unión ocurre entre el exón 4 de *SMARCA4* y el exón 2 de *CARM1*, conservando casi la secuencia entera funcional de *CARM1* y sus dominios con actividad metiltransferasa en residuos de arginina (Figura 22A). Las funciones de metilación de *CARM1* son esenciales para la leucemogénesis y en el arresto de la diferenciación de CTH y progenitores mieloides ([Greenblatt et al., 2018](#)).

TPM4::KLF2 surge de una deleción intracromosómica (situados a una distancia de ~260 kb) entre el exón 1 de la tropomiosina alfa-4 y el exón 3 del factor 2 similar a Krüpel, factor de transcripción de tipo dedo de zinc. El producto conserva los motivos de unión a DNA tipo C2H2 de KLF, localizados en el extremo carboxilo terminal (en 3') (Figura 22B). Sin embargo, aunque se ha reportado frecuentemente en LMA, leucemia promielocítica aguda, y en un par de casos de LLA-B, no se conoce su función ni se ha considerado como una posible alteración conductora (Locher et al., 2017; Jann et al., 2020).

A

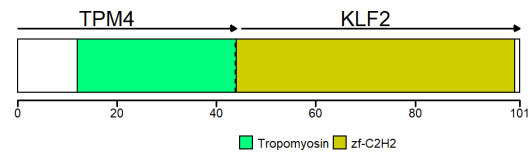
Caso #528



QLQ: Glutamine-Leucine-Glutamine; CARM1: Histone-arginine methyltransferase CARM1 N-terminal; PRMT5: SAM-dependent arginine-N-methyltransferase PRMT-type; TAD: Transactivation domain

B

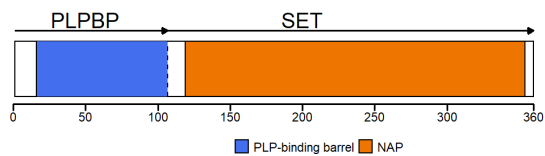
Caso #203



ZF-C2H2: Zinc finger C2H2-type

C

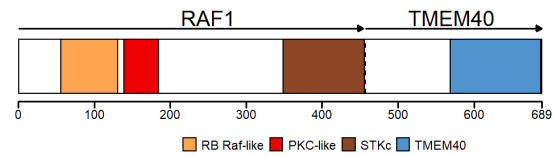
Caso # 475



NAP: Nucleosome-assembly protein

D

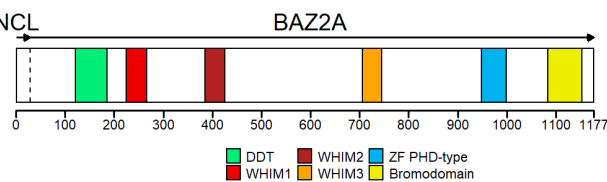
Caso #540



RB Raf-like: RAF-like RAS binding; PKC-like: Protein kinase C-like; STKc: Catalytic domain of the Serine/Threonine Kinase; TMEM40: Transmembrane protein 40 family

NCL

Caso # 475



DDT: Domain in different transcription and chromosome remodeling factors

Figura 22. Fusiones sin subtipo asignado en LLA. Las fusiones *SMARCA4::CARM1* (A) y *RAF1::TMEM40* se han reportado en tipos de cáncer no hematológico. *PLPBP::SET* y *NCL::BAZ2A* (C) se encontraron en el mismo paciente y no han sido previamente identificadas, mientras que *TPM4::KLF2* (D) se ha reportado en leucemia promielocítica aguda.

En un caso se identificaron las fusiones *PLPBP::SET* y *NCL::BAZ2A*, no reportadas previamente. El proto-oncogén *SET* conserva su único dominio (NAP) en esta fusión (Figura 22C), implicado en múltiples funciones de remodelación de la cromatina, principalmente como inhibidor de acetilación de histonas. Se han identificado varios genes compañeros que involucran a *SET* como gen en posición 5' o en 3' en distintos tipos de leucemias agudas y de diferentes linajes. La fusión *NCL::BAZ2A* del mismo paciente une a partir del exón 14 de la proteína bromodominio 2A adjacente al dominio dedo de zinc BAZ2A con el exón 1 y parte del exón 2 del gen de la nucleolina *NCL* (Figura 22C). En la fusión predicha se conservan la mayoría de los dominios tipo bromodominio con actividad de metiltransferasas de DNA, encargadas de la formación de heterocromatina y silenciamiento de genes de BAZ2A. Aunque no se conoce el papel de *PLPBP::SET* y *NCL::BAZ2A* por ser eventos nuevos, tanto *SET* y *BAZ2A* están implicados en funciones de silenciamiento transcripcional.

Características clínicas y pronóstico de los subtipos moleculares

Los subtipos identificados presentaron características clínicas que concuerdan con lo reportado en esta enfermedad (Tabla 2). Todos los casos con la fusión *ETV6::RUNX1* presentaron características de buen pronóstico (mediana de edad de 3.6 años y CL $<50 \times 10^9/L$). Anteriormente, la fusión recurrente *TCF3::PBX1* era considerada de mal pronóstico por su propensión a generar recaídas en SNC, situación que ha mejorado substancialmente con intensificación de quimioterapia (en ocasiones se considera aún de riesgo intermedio). En este estudio, los dos casos *TCF3::PBX1* tuvieron una evolución favorable (sin RT o muerte) a pesar de contar con características de RA (ambos mayores de 10 años, uno de ellos con hiperleucocitosis). Los subtipos con actividad de cinasas como *BCR::ABL1* o *CRLF2* (Ph-like en general) presentaron alta CL al diagnóstico (mediana de CL: $136.4 \times 10^9/L$ para *BCR::ABL1* y $61.7 \times 10^9/L$ para *CRLF2-r*) y una mediana de edad superior al resto de los subtipos (mediana *CRLF2*= 13 años, *BCR::ABL1*= 12 años). Solo un caso *BCR::ABL1+* y uno con *CRLF2-r* tuvieron menos de 10 años al diagnóstico (8.5 y 7

años, respectivamente). Interesantemente, el único paciente *BCR::ABL1*+ menor de 10 años presentó la isoforma mayor de *BCR::ABL1*, muy poco frecuente en LLA. Dentro de los subtipos “Otros B”, la mediana de edad de *DUX4* y *MEF2D* fue de 11.6 años, mientras que la CL suele ser menor que la de los subtipos con actividad de cinasa (mediana de CL de *MEF2D*= 61.2 x10⁹/L y *DUX4*= 29.6 x10⁹/L).

El resto de las alteraciones consideradas dentro del grupo de “Otros B” estuvieron presentes en menos de 3 pacientes, por lo que es difícil encontrar patrones de características clínicas asociadas. Ambos casos con fusiones de *PAX5* presentaron alta CL al diagnóstico (81 y 143 x10⁹/L), pero marcada diferencia en la edad de presentación de la enfermedad (1.6 y 10.4 años), mientras que el único caso con fusiones de *ZNF384* tuvo características clínicas de buen pronóstico (2 años, 2x10⁹/L de CL).

Con excepción de *ETV6::RUNX1* y *TCF3::PBX1*, el pronóstico del resto de los grupos identificados fue en general desfavorable. De estos dos subtipos, solo un caso *ETV6::RUNX1*+ recayó, y ninguno con estas alteraciones falleció durante el seguimiento. Setenta por ciento de los casos con algún subtipo dominado por cinasas, *BCR::ABL1* y *CRLF2*, presentaron algún evento durante el seguimiento, sea RT o muerte (Figura 23B y Tabla 3). La SG a 4 años de estos dos subtipos fue del 50% (Figura 23A y Tabla 3). En particular, los casos con la fusión *P2RY8::CRLF2* tuvieron el peor pronóstico: la mitad de ellos presentaron RT y solo 1 de 4 casos con esta fusión permaneció con vida durante el período de seguimiento. Esto es de gran relevancia, debido a que un caso con esta fusión presentó características clínicas de buen pronóstico y se le administró un tratamiento para RH (el paciente falleció por sepsis y actividad leucémica). Por su frecuencia y proclividad a recaer, *P2RY8::CRLF2* puede considerarse como un marcador de mal pronóstico y beneficiarse de un tratamiento más intenso aun cuando algunos casos con esta fusión presenten características clínicas de buen pronóstico.

La SG y SLE (33% de ambas) de los pacientes *DUX4* observada en este estudio contrasta con lo reportado en cohortes caucásicas, en donde se considera un marcador de excelente pronóstico (Tabla 3). Los cuatro casos con alteraciones de *DUX4* que recayeron también fallecieron; las recaídas tardaron más

tiempo en ocurrir que las de otros subtipos como *CRLF2* (promedio de RT para *CRLF2*= 6.28 meses, promedio de RT para *DUX4*= 22.3 meses) (Figura 23D). La única variable clínica asociada a RT en *DUX4* fue la edad, ya que todos los que recayeron tenían más de 10 años (tres de ellos presentaron una baja CL al diagnóstico).

Dentro del resto de los subtipos “Otros B”, dos casos de los 3 con fusiones *MEF2D* evolucionaron favorablemente pese a contar con características de RA. Ninguno de los casos con la fusión *MEF2D::SS18* recayó o murió durante el período de seguimiento, mientras que el paciente con la fusión *MEF2D::HNRNPUL1* presentó recaída muy temprana (durante los primeros 18 meses) y falleció. Hasta ahora no se conoce la relevancia biológica del gen en posición 3’ con el que se une a una *MEF2D* (con excepción de *CSF1R* que es Ph-like), ya que no se conservan ninguno de los dominios funcionales de ambos genes compañeros (Figura 14). Por otra parte, los dos casos con fusiones conductoras de *PAX5* tuvieron una evolución clínica completamente diferente. El caso con la fusión *PAX5::NOL4L*, marcador recurrente en LLA-B, no presentó RT o muerte a pesar de presentar hiperleucocitosis al diagnóstico, mientras que el paciente con la fusión nueva *PAX5::ST18* tuvo fallo a la inducción y murió durante la primera semana de tratamiento. En los casos *PAX5-r*, la única variable clínica asociada fue la gran diferencia de edad entre ambos (1.6 años el caso *PAX5::NOL4L* y 10.4 años el caso *PAX5::ST18*).

El grupo de pacientes en los que se identificaron fusiones sin subtipo asignado en LLA y con significancia pronóstica desconocida también tuvieron un curso clínico desfavorable (Figura 23E, 23F y Tabla 3). Todos estos pacientes fueron diagnosticados dentro de un rango de edad de buen pronóstico (entre 1.5 a 5.6 años) y solo el caso con la fusión *PLPBP::SET* presentó elevada CL al diagnóstico (Figura 22C). De hecho, este caso tuvo la segunda mayor CL en sangre periférica de todos los pacientes estudiados ($521 \times 10^9/L$). Tanto la SG como la SLE de este grupo de pacientes fue de apenas 25% a 4 años de seguimiento (Tabla 3).

Dos subtipos solo se presentaron en un solo paciente: el caso con la fusión *TCF3::ZNF384*, el cual recayó apenas a los 6 meses y falleció un mes después de haber recaído, y el paciente con la fusión *CEBPE::IGH*, el cual tuvo una buena evolución clínica (sin RT o muerte) a pesar de su edad como variable pronóstica adversa (14 años al diagnóstico).

Finalmente, en los pacientes en los que no se identificó ninguna fusión génica, también se observó una baja supervivencia (Figura 23G, 23H y Tabla 3). Casi la mitad de los pacientes LLA-B sin fusiones identificadas fallecieron; de ellos, 57% murieron dentro de los primeros 6 meses post-diagnóstico y 85% dentro de los primeros 18 meses. También se pudo identificar que un porcentaje importante de pacientes en este grupo contaban con características clínicas de buen pronóstico; 40% de ellos menores de 10 años y 67% de ellos con una CL menor a $50 \times 10^9/L$. Los dos casos con LLA-T en los que no se identificó ninguna fusión siguieron vivos durante el seguimiento de este estudio.

Tabla 2. Variables clínicas asociadas a subtipos genéticos en LLA

Subtipo molecular	Cuenta leucocitaria ($\times \text{mm}^3$)		<i>p</i> *	Edad (meses)		<i>p</i> *
	Mediana	Rango		Mediana	Rango	
<i>BCR::ABL1</i> (Ph)	136,400	16,956 – 641,000	0.07	144	104 - 208	0.1
<i>CRLF2</i>	61,780	10,860 – 166,000		156	86 - 202	
<i>DUX4</i>	29,590	2,920 – 138,900		140	71 - 173	
<i>ETV6::RUNX1</i>	34,800	8,990 – 41,620		44	39 - 88	
<i>MEF2D</i>	61,250	11,850 – 61,600		140	33 - 140	
Sin fusiones detectadas	13,300	810 – 167,180		129	32 - 175	

Fueron excluidos los subtipos *PAX5-r*, *TCF3::PBX1*, *IGH::CEBPE*, *TCF3::ZNF384* y fusiones no reportadas por estar presentes en 2 o menos casos de cada fusión/subtipo.

* Valor de *p* con prueba de Kruskal-Wallis entre *BCR::ABL1*, *CRLF2*, *DUX4* y grupos restantes

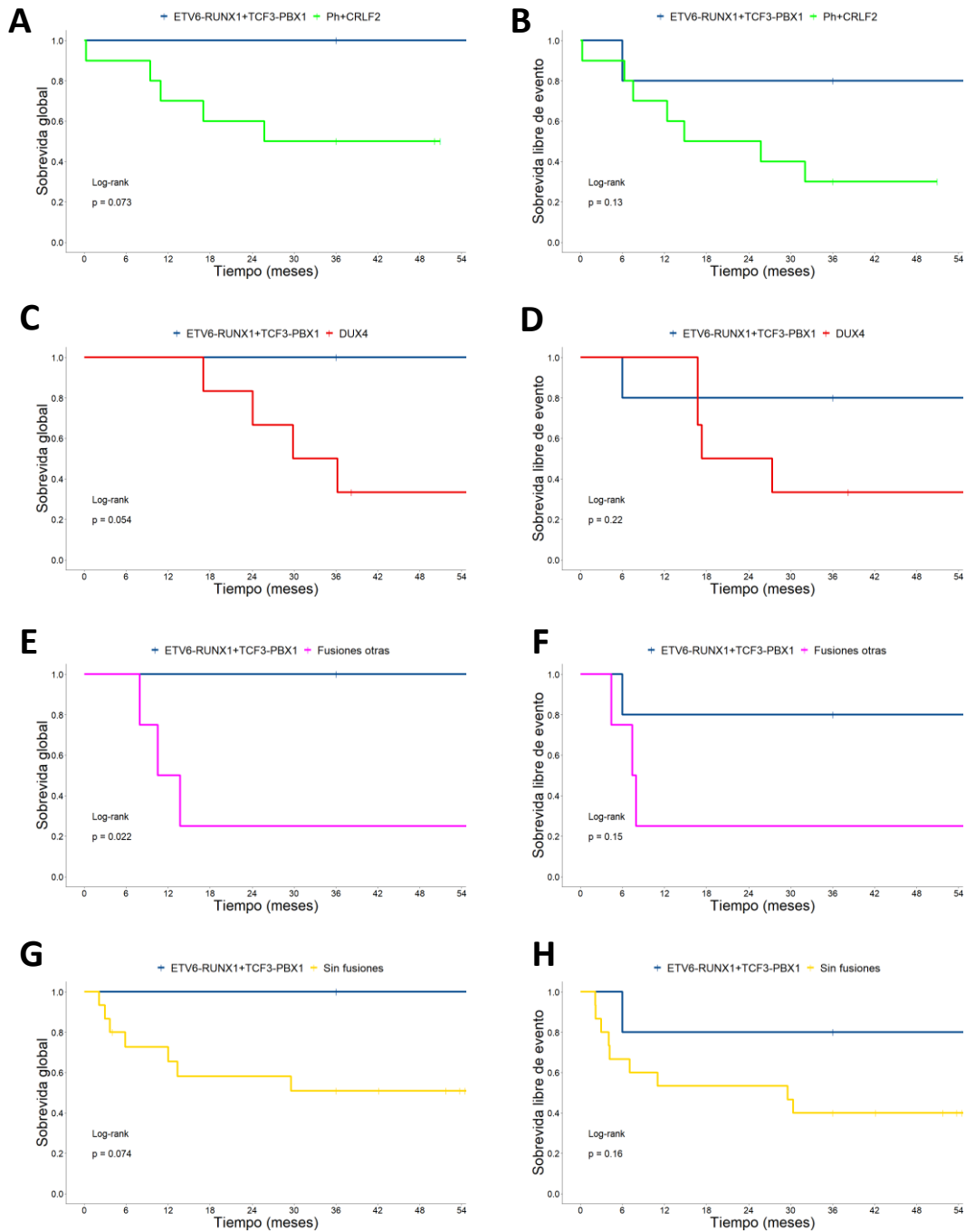


Figura 23. Probabilidad de sobrevivida global y sobrevivida libre de evento de subtipos genéticos de LLA. *ETV6::RUNX1* y *TCF3::PBX1* fueron usados de referencia en un solo grupo por considerarse de buen pronóstico (solo una recaída y ninguna muerte). Se muestra el valor de p de prueba Kaplan Meier para la probabilidad de sobrevivida global y sobrevivida libre de evento de los subtipos con actividad de cinasa *BCR::ABL1* (Ph) y *CRLF2* (A y B), *DUX4* (C y D), fusiones sin subtipo asignado (E y F, ver figura 22) y pacientes sin fusiones identificadas (G y H). Se excluyeron los subtipos *MEF2D*, *PAX5* y *ZNF384* en este análisis.

Tabla 3. Sobrevida de subtipos genéticos en LLA

Subtipo molecular	SG a 4 años	SLE a 4 años
Buen pronóstico*	100 ± 0	80 ± 18
Ph + CRLF2	50 ± 16	30 ± 14
DUX4	33 ± 19	33 ± 19
Otros B **	50 ± 20	33 ± 19
Fusiones sin subtipo asignado***	25 ± 21	25 ± 21
Sin fusiones detectadas	51 ± 13	40 ± 12

* Buen pronóstico incluye a *ETV6::RUNX1* y *TCF3::PBX1* en este estudio

** En Otros B se incluyen subtipos *MEF2D*, *PAX5* y *ZNF384*

*** Fusiones sin subtipo conocido en LLA y/o no reportadas previamente en la literatura

SG= Sobrevida global; SLE= Sobrevida libre de evento

DISCUSIÓN

La administración de esquemas de quimioterapia ajustados en función del riesgo a recaída y muerte son un paradigma en el tratamiento contra el cáncer, y han permitido que casi el 99% de los niños con LLA en países de alto ingreso alcancen la remisión y la mayoría de ellos pueden ser curados. El conocimiento generado en décadas recientes ha permitido, entre otras cosas, identificar nuevos subtipos moleculares con pronóstico asociado y monitorear la presencia de EMR como respuesta al tratamiento. A falta de variables pronósticas que impidan modular la intensidad del tratamiento, se dificulta asignar el esquema óptimo que evite toxicidad innecesaria o intensidad insuficiente. Las bajas tasas de sobrevida observadas en países de bajo y mediano ingreso son, al menos en parte, consecuencia de las altas tasas de mortalidad durante la inducción y de las recaídas, cuyo pronóstico suele ser muy desfavorable ([Abdelmabood et al., 2020](#)). El uso de herramientas genómicas en décadas recientes permitió identificar nuevas alteraciones genéticas y nuevos subtipos moleculares principalmente en pacientes que no cuentan con los marcadores clásicos en LLA, los cuales además suelen estar asociados a un peor pronóstico. Estas alteraciones genéticas recién descubiertas no son fácilmente detectables con herramientas convencionales de diagnóstico molecular que no pueden cubrir la compleja heterogeneidad de la enfermedad. La secuenciación del transcriptoma (RNA-seq) ha sido una de las herramientas más usadas en estudios genómicos en LLA, capaz de identificar múltiples alteraciones con relevancia pronóstica. Las fusiones génicas están entre las alteraciones más comunes y son eventos conductores en la leucemogénesis, además de dictar el comportamiento biológico, clínico, e incluso ser blanco de terapia dirigida. Esta investigación constituye el primer análisis de secuenciación del transcriptoma dirigido a la identificación de fusiones génicas/transcritos quiméricos asociados a recaída temprana y muerte en LLA pediátrica en México.

Una de las principales observaciones que sugieren la presencia de fusiones génicas aún no descritas en México es la alta proporción de casos con características clínicas de RA. Aproximadamente un tercio de

los niños con LLA se clasifican como de alto riesgo en cohortes estadounidenses y europeas (Zaliova et al., 2019), mientras que en México entre el 45-70% tienen criterios de alto riesgo cuando son diagnosticados (Schultz et al., 2007; Núñez-Enríquez et al., 2019; Muñoz-Aguirre et al., 2021). La edad es la principal variable asociada a este fenómeno. Alrededor de 30% los niños con LLA en México tienen más de 10 años al diagnóstico, similar a lo reportado en población hispana en EE. UU. pero mayor que en poblaciones caucásicas (23-25%) (Kahn et al., 2017; Martín-Trejo et al., 2017; Jeha et al., 2019; Núñez-Enríquez et al., 2019; Zaliova et al., 2019, Muñoz-Aguirre et al., 2021; Raca et al., 2021). La incidencia de LLA en niños entre 5 y 15 años continúa siendo alta en México, cifras similares a las de Ecuador, Costa Rica, Colombia y algunas regiones de Brasil (de Souza Reis et al., 2011; Steliarova-Foucher et al., 2017; Flores-Lujano et al., 2022). En nuestro estudio se observó que la edad mayor a 10 años es la variable que más se asocia a una menor sobrevida. Sin embargo, no se observó el valor pronóstico esperado de la CL, ya que el grupo de casos con una cuenta $<50 \times 10^9/L$ también tuvieron una menor SG y SLE, a pesar de ser una variable de buen pronóstico (su probabilidad de sobrevida fue comparable a los casos con hiperleucocitosis). Por esta razón, no se encontraron diferencias en la sobrevida cuando se consideran estas variables, edad y CL, en dos grupos de riesgo (habitual y alto). Es conocido que la optimización de tratamiento es más útil para los casos de riesgo intermedio (DelRocco et al., 2021), en quienes la intensificación de este puede evitar recaídas cuando se consideran otros criterios como la EMR o clasificación genética (Schultz et al., 2007). Tanto la edad como la CL son variables independientes de pronóstico, parcialmente relacionadas con la presencia de subtipos genéticos agresivos en niños mayores, como *BCR::ABL1* o Ph-like, o de subtipos de buen pronóstico como *ETV6::RUNX1* y las hiperdiploidías presentes casi de forma exclusiva en el primer pico de incidencia de la enfermedad (~ a los 4 años). La mediana de edad de los casos *ETV6::RUNX1* en este estudio coincide con lo anterior (~3.6 años), en contraste con prácticamente el resto de los subtipos genéticos identificados en LLA con mayor edad a su presentación. Por ser una sub-cohorte orientada a identificar alteraciones genéticas de RA, es

esperable la alta frecuencia de subtipos con actividad de cinasa como *BCR::ABL1* o la presencia de alteraciones de *CRLF2*/Ph-like, que en conjunto representan 21.7% de los casos de LLA-B de este estudio. A pesar de que Ph-like se identifica con expresión génica, las alteraciones subyacentes son las que permiten subclasificarlo (Tasian et al., 2017). Debido a su prevalencia en todas las edades, mal pronóstico asociado y ser clínicamente accionable con ITC, la OMS lo reconoce como una entidad única en LLA (Alaggio et al., 2022). Aunque no existe un estándar de oro, en general hay buena concordancia entre distintas plataformas tecnológicas empleadas para identificar Ph-like, que van desde perfiles de expresión global con microarreglos o RNA-seq, hasta el uso de paneles de genes de mucho menor costo y mayor aplicabilidad clínica. Con el análisis paralelo de nuestra sub-cohorte de microarreglos pudimos comprobar que las firmas de 8 o 15 genes del Children's Oncology Group de EE. UU. pueden identificar casos Ph-like, algunos de los cuales contaban con rearrreglos de *CRLF2* identificados con RNA-seq. Esto se pudo lograr incluso considerando aspectos técnicos que pudieran dificultar su reproducibilidad, por lo que la implementación de estas firmas génicas puede realizarse con herramientas más rápidas y económicas (Chiaretti et al., 2018; Chiaretti et al., 2019), algunas de las cuales ya han comenzado a utilizarse en México y Colombia (Linares-Ballesteros et al., 2022; Martínez-Anaya et al., 2022). La identificación del perfil Ph-like requiere forzosamente de la búsqueda de la alteración subyacente, ya que el perfil de expresión *per se* no ofrece información relevante para una intervención terapéutica. Las alteraciones de *CRLF2* son las más comunes en Ph-like, el cual forma un heterodímero con IL7R para inducir proliferación de progenitores multipotentes mediante fosforilación de proteínas STAT (Reche et al., 2001; Scheeren et al., 2010). La sobre-expresión de *CRLF2* por secuestro de promotores (fusión *P2RY8::CRLF2*) o de potenciadores (*IGH::CRLF2*) es altamente sugestiva de la presencia de un rearrreglo genético involucrado, aunque una minoría de casos con expresión de *CRLF2* (<2%) no son Ph-like (Reshmi et al., 2017; Roberts et al., 2018). En este estudio la frecuencia de casos de LLA-B con alteraciones de *CRLF2* fue de 10.9%, similar a la reportada en cohortes no seleccionadas en EE. UU., que van desde 2%

hasta 15% (Roberts et al., 2014; Dou et al., 2017; Imamura et al., 2017; Reshmi et al., 2017; Schmäh et al., 2017). Las muy variables frecuencias reportadas de *CRLF2-r* suelen deberse a las características clínicas presentes o por las metodologías empleadas (Tasian et al., 2017). Sin embargo, son numerosas las evidencias de que Ph-like y especialmente las alteraciones de *CRLF2* son más frecuentes en pacientes con ascendencia latina en EE. UU. (Harvey et al., 2010; Harvey et al., 2010; Raca et al., 2021). La frecuencia de *CRLF2-r* en este estudio es similar a población de Guatemala (11%), más alta que en Colombia (6.6%), pero más baja que en otro estudio en México (38%), diferencias posiblemente debidas a que se usaron distintas herramientas (RNA-seq, citometría de flujo, o PCR) (Linares-Ballesteros et al., 2022; Juárez-Velázquez et al., 2022; Lee et al., 2022). La expresión de *CRLF2* en el ámbito clínico puede evaluarse desde la inmunofenotipificación o PCR cuantitativa para identificar los casos que ameriten análisis posteriores para identificar la alteración subyacente. No obstante, es necesario descartar los casos que pueden presentar expresión pero no una alteración conductora subyacente, ya que ocasionalmente la fusión *P2RY8::CRLF2* es un evento subclonal que co-ocurre con otros subtipos (Vesely et al., 2017; Potter et al., 2019). En dos casos *P2RY8::CRLF2* de nuestra población de estudio se identificó otro marcador concomitante, uno con *IGH::CRLF2* (en una baja cantidad de lecturas, por lo que se determinó *P2RY8::CRLF2* como subtipo primario) y un caso con la fusión nueva *CDKN2A::TRDC*. Las alteraciones somáticas de *CDKN2A*, principalmente deleciones, están presentes en el 80% y 50% de los casos de LLA-T y LLA-B, respectivamente, y son la segunda alteración más común en LLA *BCR::ABL1+* (Williams et al., 2007; Ma et al., 2018; Mullighan et al., 2008). La pérdida de *CDKN2A* (*ARF* en especial) disminuye la latencia de la enfermedad en modelos *in vivo* *BCR::ABL1+* con deleciones en *IKZF1* (Churchman et al., 2015). Consistente con esto y a pesar de tener baja CL al diagnóstico (13.7×10^9 /L), el caso con la fusión concomitante *CDKN2A::TRDC* tuvo la progresión más rápida de todos los casos *P2RY8::CRLF2* (recayó a los 6 meses y falleció a los 9 meses post-diagnóstico).

Se han identificado algunas variantes de ganancia de función en el gen *GATA3* asociadas a Ph-like y etnicidad hispana (Pérez-Andreu et al., 2011). La sobre-expresión de *GATA3* facilitaría la accesibilidad de la cromatina en regiones proximales a *CRLF2*, haciéndolo susceptible a sufrir rearrreglos génicos (Yang et al., 2022). Por esta razón, la implementación de técnicas que identifiquen Ph-like, y en especial los subtipos con alteraciones de *CRLF2*, debe considerarse en futuros análisis. La fusión *P2RY8::CRLF2* puede identificarse fácilmente con RT-PCR, y esto puede complementarse evaluando la expresión de *CRLF2* con citometría de flujo desde la inmunofenotipificación del paciente o con PCR cuantitativa. Esto también es importante debido a que *P2RY8::CRLF2* puede presentarse a menor edad (a diferencia de *IGH::CRLF2* que suele ser más común en adolescentes), y por lo tanto es menos probable que reciba tratamiento para RA (Russell et al., 2017). El 75% de los casos *P2RY8::CRLF2* de nuestro análisis fallecieron durante el seguimiento, uno de ellos con características clínicas de buen pronóstico y que recibió un tratamiento para RH. Además de la intensificación de la quimioterapia, la incorporación de inhibidores de cinasas JAK como el ruxolitinib ha mostrado eficacia en algunos reportes de caso para el tratamiento de pacientes Ph-like con alteraciones de *JAK2* y *EPOR* (Ding et al., 2018; Niswander et al., 2021; Chen et al., 2022). Su potencial es menos claro para *CRLF2-r* fuera de ensayos pre-clínicos (Zhang et al., 2018), a pesar de que las mutaciones en genes *JAK* y la activación constitutiva de la vía JAK-STAT es distintiva de este subtipo (Reshmi et al., 2017). La eficacia del ruxolitinib está siendo explorada en ensayos clínicos de fase III, aunado a la búsqueda de nuevas estrategias terapéuticas para la inhibición de JAK-STAT y Ph-like *CRLF2-r* (Aldoss y Advani, 2021; Chang et al., 2021; Sasaki et al., 2022).

Uno de los hallazgos más notables en nuestro estudio es la baja sobrevida de los casos *DUX4*, que normalmente tiene un excelente pronóstico en cohortes estadounidenses, europeas y asiáticas (Novakova et al., 2021; Li et al., 2021; Brady et al., 2022). La sobrevida global a 4 años de nuestros casos *DUX4* fue de apenas 33%, en contraste con más de 95% en EE. UU. (Gu et al., 2019; Brady et al., 2022). El buen pronóstico de este subtipo se ha logrado con intensificación del tratamiento, ajustado debido a que

en *DUX4* es frecuente la persistencia de EMR. Al menos la mitad de los casos *DUX4* suelen presentar elevada EMR durante las primeras semanas del tratamiento (Gu et al., 2019; Li et al., 2021). En el protocolo St Jude XVI, todos los pacientes *DUX4* recibieron tratamiento estándar (riesgo intermedio) con intensificación durante la fase de consolidación, por persistencia de EMR al día 8 en sangre periférica (Jeha et al., 2021). Ningún paciente *DUX4* recayó bajo el protocolo St Jude XVI (Jeha et al., 2021), lo que demuestra que la intensificación del tratamiento es clave para el buen pronóstico asociado a algunos subtipos genéticos. En el presente estudio, los 4 casos *DUX4* que presentaron recaída posteriormente fallecieron, aunque el tiempo a la recaída fue mayor (19.5 meses en promedio) en comparación de otros subtipos como *BCR::ABL1* o *CRLF2-r* (14.6 meses en promedio). La edad parece ser la única variable clínica asociada al mal pronóstico de *DUX4*, ya que los únicos dos casos que no recayeron o murieron eran menores de 10 años (uno incluso presentó elevada CL). Otra variable que puede dificultar el monitoreo de la enfermedad en subtipos como *DUX4* es su plasticidad inmunofenotípica, que lleva a los blastos a expresar frecuentemente marcadores de linajes ambiguos, principalmente mieloides (Slamova et al., 2014). Este fenómeno de cambio de linaje, también llamado cambio monocítico o monocitoide (“monocytic switch”) es frecuente en otros subtipos con arresto muy temprano de la diferenciación como *ZNF384* o *PAX5-P80R* (Wu et al., 2021), y en *DUX4* se presenta con mayor frecuencia en niños mayores (Novakova et al., 2021). El cambio de linaje puede ocurrir antes del tratamiento y generar resistencia intrínseca en poblaciones subclonales, o surgir bajo presión selectiva del tratamiento (Dobson et al., 2020). No se ha observado que este fenómeno impacte siempre en el pronóstico, pero la pérdida de CD19 (marcador de linaje B) puede subestimar la presencia de blastos residuales durante el monitoreo de EMR, o limitar la efectividad de las terapias inmunológicas dirigidas contra CD19 (Wang et al., 2017; Novakova et al., 2021; Semchenkova et al., 2022). Algunos reportes preliminares en México refieren dificultades en la distinción de blastos leucémicos en la evaluación de EMR, o una frecuencia inesperadamente alta de marcadores mieloides en LLA-B (Gonzalez-Montalvo et al., 2017; Juárez-

[Avendaño et al., 2020](#)). El pronóstico de *DUX4* también puede estar influido por la presencia de subgrupos con expresión aberrante de *FLT3* y *CEBPA* dentro del mismo subtipo, como ha sido reportado recientemente ([Brady et al., 2022](#)). Por otra parte, el temprano arresto en su diferenciación y su perfil transcripcional similar al de progenitores megacariocíticos-eritroides se asocia a fenómenos de adhesión, quimiotaxis, y señalización de citocinas/quimiocinas del eje CXCR4-CXCL12, CCL2 y CXCL8 ([Dmitriev et al., 2016](#); [de Rooij et al., 2017](#); [Tanaka et al., 2018](#); [Schroeder et al., 2019](#)). En un análisis previo de nuestro grupo de trabajo en el que se usaron microarreglos de expresión, identificamos a la interleucina 8 (CXCL8) como uno de los genes codificantes con mayor expresión en pacientes que presentaron RT ([Bárcenas-López et al., 2020](#)). CXCL8, una quimiocina proinflamatoria de origen monocítico positivamente regulada por la vía NF- κ B, favorece la adhesión de blastos leucémicos a células estromales y contribuye a mecanismos de supervivencia y quimiorresistencia ([de Vasconcellos et al., 2011](#)).

Considerando que los factores de transcripción como CEBPA actúan en la diferenciación monocítica ([Schroeder et al., 2019](#)), la plasticidad característica de *DUX4* es un factor para considerar en el monitoreo de la respuesta a la quimioterapia y la incorporación de marcadores de linaje monocítico durante la progresión de la enfermedad. Sin embargo, la identificación de alteraciones de *DUX4* es compleja incluso dentro del contexto del análisis genómico ([Lilljebjörn et al., 2016](#); [Tian et al., 2019](#)).

Algunas moléculas que han emergido como buenos marcadores inmunofenotípicos que sugieren la presencia de alteraciones de *DUX4* son CD2 y CD371 (codificada por el gen *CLEC12A*), siendo este último posible blanco transcripcional de *DUX4* ([Slamova et al., 2014](#); [Schinnerl et al., 2019](#)).

A pesar de haber podido clasificar alrededor del 63% de nuestros casos de LLA-B mediante la presencia de fusiones conductoras o expresión génica aberrante, 6.5% de los casos presentaron fusiones que no se pudieron asignar a ningún subtipo conocido en LLA, y en 32.6% no se encontraron alteraciones. Un patrón en común en las fusiones sin subtipo asignado es a presencia de genes involucrados en funciones epigenéticas (9% de todas las fusiones involucran al menos un regulador epigenético). Una de ellas,

SMARCA4::CARM1, ha sido reportada en distintos tipos de adenocarcinomas de vejiga, cérvix, pulmón, ovario, y mama, y glioblastoma multiforme (Kangaspeska et al., 2012; Gao et al., 2018). *SMARCA4* es un conocido supresor tumoral recurrentemente mutado en tumores pediátricos (Huether et al., 2014; Gröbner et al., 2018), mientras que su gen parálogo *SMARCA2* se ha identificado en fusiones en LLA, pero no como la alteración conductora (Li et al., 2018; Gu et al., 2019). La alta expresión de *CARM1* se ha asociado a mal pronóstico en distintos tipos de cáncer de origen epitelial; sin embargo, las amplificaciones y variantes de un solo nucleótido (SNVs) son las alteraciones más comunes en *CARM1*, mientras que las fusiones son raras (Wang et al., 2014; Qiu y Wang et al., 2022). Un hallazgo interesante en el paciente con la fusión *SMARCA4::CARM1* fue la presencia de otras fusiones que, aunque fueron descartadas por nuestros filtros de selección, están involucrados genes de la misma vía. Genes como *ARID1B*, *BCL11A*, además de *SMARCA4* y *CARM1* pertenecen directa o indirectamente a la familia SWI-SNF. SWItch/Sucrose Non-Fermentable SWI-SNF es un complejo proteico de remodeladores de cromatina dependiente de ATP, encargado de regular la transcripción mediante el desplazamiento o desalojo de nucleosomas (Mittal y Roberts, 2020). Estos mecanismos de activación transcripcional, en especial aquellos que involucran factores de transcripción maestros, son un sello distintivo en leucemias agudas con baja carga mutacional y menor latencia para su desarrollo como las del subtipo *KMT2A* en niños menores de 1 año (Neff y Armstrong, 2013; Huether et al., 2014). De hecho, los pacientes en los que se detectaron fusiones que involucran genes con función epigenética como *SMARCA4::CARM1* y el caso con las fusiones *PLPBP::SET* y *NCL::BAZZA*, tenían menor edad a su diagnóstico (mediana= 4 años), una SG y SLE muy pobre (ambas 25%, la más baja en conjunto en este análisis) y rápida progresión de la enfermedad (en menos de 9 meses presentaron recaída y/o muerte). Solo uno de ellos presentó hiperleucocitosis, ($521 \times 10^9/L$), la cual es una característica clínica común también en los rearrreglos de *KMT2A* (Kong et al., 2014). Conocer el significado clínico de estos subtipos raros es muy desafiante, ya que constituyen una muy baja proporción de todos los casos en LLA. El empleo de nuevos abordajes

terapéuticos mutación-agnósticos como la inmunoterapia o inhibidores de checkpoints inmunológicos son algunas estrategias para aquellos subtipos cuyos mecanismos biológicos son desconocidos (Buscarlet et al., 2014; Daver N, 2021; Qiu y Wang et al., 2022; Sheykhhasan et al., 2022).

La fortaleza de este estudio es el análisis de transcriptoma completo que permite identificar alteraciones genéticas a escala global que no es del todo posible cuando se usan paneles de genes, especialmente por la posibilidad de encontrar alteraciones no conocidas previamente. En esta investigación, nos enfocamos en analizar casos con pronóstico desfavorable y/o características clínicas de alto riesgo que identificamos de forma retrospectiva, asociando la presencia de subtipos genéticos con la progresión de la enfermedad. La principal limitación de este proyecto es el tamaño de muestra, que nos impide obtener conclusiones sobre el porcentaje real de los subtipos moleculares identificados dada la compleja heterogeneidad que caracteriza a la LLA. Sin embargo, el hecho de que distintos estudios en México hayan identificado como un fenómeno común elevado porcentaje de casos de LLA pediátrica con características de alto riesgo nos sugieren focalizar nuestros estudios sobre los casos con peor pronóstico. Diseñar análisis similares con un mayor número de individuos y en cohortes no seleccionadas (o consecutivas) nos permitirá conocer mejor las frecuencias de cada subtipo identificado (y otros no identificados en esta investigación) e inferir conclusiones apropiadas respecto al pronóstico clínico asociado a ellos. Para subtipos como *MEF2D*, *PAX5* y *ZNF384*, que en general se consideran de pronóstico intermedio o malo, no fue posible inferir su valor pronóstico en nuestra población. A pesar de esto, identificar estos subtipos puede ofrecer oportunidades terapéuticas en un futuro, como muestran estudios pre-clínicos sobre el uso de inhibidores de desacetilasas de histonas o inhibidores de cinasas de señalización BCR para *MEF2D*-r (Gu et al., 2016; Tsuzuki et al., 2020; Zhang et al., 2022), e inhibidores de FLT3 para *ZNF384*-r (Dickerson et al., 2022; Zhao et al., 2022). Aunque la secuenciación de RNA es una herramienta muy versátil que incluso permite identificar SNVs (Gu et al., 2019), duplicaciones internas en tándem (ITD) (Tian et al., 2020), e incluso aneuploidías y alteraciones en el número de copias (CNAs)

(Bařinka et al., 2022), su análisis todavía es complejo y poco práctico para su aplicación clínica. Algunas alteraciones son difíciles de detectar incluso con RNA-seq, como las que involucran el locus *IGH*. La profundidad y longitud de las lecturas (preferentemente ≥ 75 bases), y el protocolo de elaboración de bibliotecas (depleción de RNA ribosomal vs selección de RNAs poliadenilados) son aspectos técnicos que pueden limitar la sensibilidad de algunos algoritmos para identificar transcritos quiméricos de baja expresión y/o difíciles de mapear al genoma de referencia (Tian et al., 2020). El uso de otras herramientas ómicas complementarias será muy útil en futuros esfuerzos para identificar alteraciones conductoras en casos que aún evaden la clasificación genética.

Como se ha mencionado, algunos marcadores inmunofenotípicos pueden sugerir la presencia de ciertos subtipos genéticos e identificarse con herramientas más económicas como la citometría de flujo, como la expresión de *CRLF2*, CD2 y CD371 para *DUX4*, expresión de marcadores mieloides en *ZNF384-r*, alta expresión de CD38 característica de *MEF2D-r* (Ohki et al., 2019), o índice de DNA para aneuploidías.

Es absolutamente necesaria la implementación del monitoreo de la EMR como variable clínica de rutina (O'Connor et al., 2018; Enshaei et al., 2020; Loh et al., 2020). La clasificación genética puede ser complementaria a la evaluación de la respuesta a la enfermedad, y muy útil cuando se usan ambos factores en la estratificación de riesgo (Jeha et al., 2021). En este estudio se observó que algunos pacientes con características clínicas de buen pronóstico se habrían beneficiado de quimioterapia más intensa. Además, la plasticidad fenotípica debe considerarse como un fenómeno que podría subestimar la proporción de blastos leucémicos a lo largo del tratamiento; todo lo anterior, con el fin de evitar las recaídas.

El pronóstico para los pacientes con LLA después de una recaída sigue siendo malo en cualquier región del mundo, y aproximadamente la mitad de los casos sobreviven más de 5 años después de la primera recaída en países de alto ingreso (Oskarsson et al., 2015; Rheingold et al., 2019). En esta sub-cohorte de casos con características clínicas de RA, menos del 20% de los casos sobrevivieron después de 4 años.

Aunque la mortalidad durante el primer año de tratamiento aún es alta en México (13.8%) (Núñez-Enríquez et al., 2019), la recaída sigue siendo uno de los mayores obstáculos para mejorar la sobrevida en LLA. Se ha demostrado que el ajuste de la intensidad de la terapia mediante el monitoreo de la enfermedad residual puede incluso atenuar las disparidades étnicas y sociales (Yang et al., 2011; Aguilar et al., 2021; Romo et al., 2021; Lee et al., 2022).

CONCLUSIONES

Con análisis de transcriptoma se pudieron clasificar más del 60% de los casos en este estudio, con el que pudimos conocer alteraciones genéticas con impacto pronóstico en niños mexicanos con LLA. Aunque la implementación de estas tecnologías con fines clínicos es aún difícil en países como el nuestro, se identificaron algunos subtipos moleculares en LLA que merecen ser considerados como parte de la clasificación genética de rutina. Los esfuerzos colaborativos son cruciales para evaluar el significado pronóstico y la frecuencia de las alteraciones genéticas en nuestra población. Además, nos permitirá identificar otros subtipos moleculares menos frecuentes y con relevancia clínica, complementar con el uso de otras herramientas ómicas e integrar el monitoreo de EMR de manera rutinaria, con el fin de proveer en un futuro una clasificación de riesgo más precisa y mejorar la sobrevida.

REFERENCIAS BIBLIOGRÁFICAS

- Abdelmabood S, Fouda AE, Boujettif F, Mansour A. Treatment outcomes of children with acute lymphoblastic leukemia in a middle-income developing country: high mortalities, early relapses, and poor survival. *J Pediatr (Rio J)*. 2020;96(1):108-116. doi:10.1016/j.jpmed.2018.07.013
- Aguilar DV, Gonzalez-Montalvo P, Romo H, et al., Molecular and MRD-Based Characterization of Acute Lymphoblastic Leukemia in Mexico: Experience from the Mexico in Alliance with St. Jude “Bridge Project”. *Blood*. 2021. 138(1); 1209. doi.org/10.1182/blood-2021-154168.
- Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms [published online ahead of print, 2022 Jun 22]. *Leukemia*. 2022;1-29. doi:10.1038/s41375-022-01620-2
- Aldoss I, Advani AS. Have any strategies in Ph-like ALL been shown to be effective?. *Best Pract Res Clin Haematol*. 2021;34(1):101242. doi:10.1016/j.beha.2021.101242
- Alexander TB, Gu Z, Iacobucci I, et al. The genetic basis and cell of origin of mixed phenotype acute leukaemia. *Nature*. 2018;562(7727):373-379. doi:10.1038/s41586-018-0436-0
- Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet*. 2018;391(10125):1023-1075. doi:10.1016/S0140-6736(17)33326-3
- Andersson AK, Ma J, Wang J, et al. The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nat Genet*. 2015;47(4):330-337. doi:10.1038/ng.3230.
- Bárcenas-López DA, Núñez-Enríquez JC, Hidalgo-Miranda A, et al. Transcriptome Analysis Identifies LINC00152 as a Biomarker of Early Relapse and Mortality in Acute Lymphoblastic Leukemia. *Genes (Basel)*. 2020;11(3):302. Published 2020 Mar 13. doi:10.3390/genes11030302
- Bařinka J, Hu Z, Wang L, et al. RNAseqCNV: analysis of large-scale copy number variations from RNA-seq data. *Leukemia*. 2022;36(6):1492-1498. doi:10.1038/s41375-022-01547-8
- Bartley K, Metayer C, Selvin S, Ducore J, Buffler P. Diagnostic X-rays and risk of childhood leukaemia. *Int J Epidemiol*. 2010;39(6):1628-1637. doi:10.1093/ije/dyq162
- Bateman CM, Colman SM, Chaplin T, et al. Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. *Blood*. 2010;115(17):3553-3558. doi:10.1182/blood-2009-10-251413
- Bekker-Méndez VC, Miranda-Peralta E, Núñez-Enríquez JC, et al. Prevalence of gene rearrangements in Mexican children with acute lymphoblastic leukemia: a population study-report from the Mexican Interinstitutional Group for the identification of the causes of childhood leukemia. *Biomed Res Int*. 2014;2014:210560. doi:10.1155/2014/210560

Bidwell JP, Torrungruang K, Alvarez M, Rhodes SJ, Shah R, Jones DR, et al. Involvement of the nuclear matrix in the control of skeletal genes: the NMP1 (YY1), NMP2 (Cbfa1), and NMP4 (Nmp4/CIZ) transcription factors. *Crit Rev Eukaryot Gene Expr* (2001) 11:279–97. doi: 10.1615/CritRevEukarGeneExpr.v11.i4

Boer JM, Marchante JR, Evans WE, et al. BCR-ABL1-like cases in pediatric acute lymphoblastic leukemia: a comparison between DCOG/Erasmus MC and COG/St. Jude signatures. *Haematologica*. 2015;100(9):e354-e357. doi:10.3324/haematol.2015.124941

Bonaventure A, Harewood R, Stiller CA, et al. Worldwide comparison of survival from childhood leukaemia for 1995-2009, by subtype, age, and sex (CONCORD-2): a population-based study of individual data for 89 828 children from 198 registries in 53 countries [published correction appears in *Lancet Haematol*. 2017 May;4(5):e201]. *Lancet Haematol*. 2017;4(5):e202-e217. doi:10.1016/S2352-3026(17)30052-2

Brady SW, Roberts KG, Gu Z, et al. The genomic landscape of pediatric acute lymphoblastic leukemia [published online ahead of print, 2022 Sep 1]. *Nat Genet*. 2022;10.1038/s41588-022-01159-z. doi:10.1038/s41588-022-01159-z

Buchmann S, Schrappe M, Baruchel A, et al. Remission, treatment failure, and relapse in pediatric ALL: an international consensus of the Ponte-di-Legno Consortium. *Blood*. 2022;139(12):1785-1793. doi:10.1182/blood.2021012328

Buscarlet M, Krasteva V, Ho L, et al. Essential role of BRG, the ATPase subunit of BAF chromatin remodeling complexes, in leukemia maintenance. *Blood*. 2014;123(11):1720-1728. doi:10.1182/blood-2013-02-483495

Carranza C, Granados L, Morales O, et al. Frequency of the ETV6-RUNX1, BCR-ABL1, TCF3-PBX1, and MLL-AFF1 fusion genes in Guatemalan pediatric acute lymphoblastic leukemia patients and their ethnic associations. *Cancer Genet*. 2013;206(6):227-232. doi:10.1016/j.cancergen.2013.05.017

Chang Y, Min J, Jarusiewicz JA, et al. Degradation of Janus kinases in CRLF2-rearranged acute lymphoblastic leukemia. *Blood*. 2021;138(23):2313-2326. doi:10.1182/blood.2020006846

Chen X, Yuan L, Zhou J, et al. Sustained remission after ruxolitinib and chimeric antigen receptor T-cell therapy bridged to a second allogeneic hematopoietic stem cell transplantation for relapsed Philadelphia chromosome-like B-cell precursor acute lymphoblastic leukemia with novel NPHP3-JAK2 fusion. *Genes Chromosomes Cancer*. 2022;61(1):55-58. doi:10.1002/gcc.22995

Chiaretti S, Akram T, Martina C, et al. The Validation of the BCR/ABL1-like Predictor across Laboratories Shows Reproducibility of Results. *Blood*. 2019;134(1):5211. doi.org/10.1182/blood-2019-122116.

Chiaretti S, Messina M, Grammatico S, et al. Rapid identification of BCR/ABL1-like acute lymphoblastic leukaemia patients using a predictive statistical model based on quantitative real time-polymerase chain reaction: clinical, prognostic and therapeutic implications. *Br J Haematol*. 2018;181(5):642-652. doi:10.1111/bjh.15251

- Choi SH, Gearhart MD, Cui Z, et al. DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. *Nucleic Acids Res.* 2016;44(11):5161-5173. doi:10.1093/nar/gkw141
- Chomczynski P, Sacchi N. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc.* 2006;1(2):581-585. doi:10.1038/nprot.2006.83
- Clark JM. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* 1988;16(20):9677-9686. doi:10.1093/nar/16.20.9677
- Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. *Nat Immunol.* 2007;8(5):463-470. doi:10.1038/ni1454
- Cogliano VJ, Baan R, Straif K, et al. Preventable exposures associated with human cancers. *J Natl Cancer Inst.* 2011;103(24):1827-1839. doi:10.1093/jnci/djr483
- Coustan-Smith E, Mullighan CG, Onciu M, et al. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. *Lancet Oncol.* 2009;10(2):147-156. doi:10.1016/S1473-2045(08)70314-0
- Cunningham F, Allen JE, Allen J, et al. Ensembl 2022. *Nucleic Acids Res.* 2022;50(D1): D988-D995. doi:10.1093/nar/gkab1049
- Daver N. Immune checkpoint inhibitors in acute myeloid leukemia. *Best Pract Res Clin Haematol.* 2021;34(1):101247. doi:10.1016/j.beha.2021.101247
- De Iaco A, Planet E, Coluccio A, Verp S, Duc J, Trono D. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet.* 2017;49(6):941-945. doi:10.1038/ng.3858
- de Rooij B, Polak R, van den Berk LCJ, Stalpers F, Pieters R, den Boer ML. Acute lymphoblastic leukemia cells create a leukemic niche without affecting the CXCR4/CXCL12 axis. *Haematologica.* 2017;102(10):e389-e393. doi:10.3324/haematol.2016.159517
- de Souza Reis R Sr, de Camargo B, de Oliveira Santos M, de Oliveira JM, Azevedo Silva F, Pombo-de-Oliveira MS. Childhood leukemia incidence in Brazil according to different geographical regions. *Pediatr Blood Cancer.* 2011;56(1):58-64. doi:10.1002/pbc.22736
- de Vasconcellos JF, Laranjeira AB, Zanchin NI, et al. Increased CCL2 and IL-8 in the bone marrow microenvironment in acute lymphoblastic leukemia. *Pediatr Blood Cancer.* 2011;56(4):568-577. doi:10.1002/pbc.22941
- DelRocco N, Loh ML, Borowitz M, et al. Comparison of Current and Enhanced Risk Stratification of 21,199 Children, Adolescents, and Young Adults with Acute Lymphoblastic Leukemia Using Objective Risk Categorization Criteria: A Children's Oncology Group Report. *Blood.* 2021; 138(1): 2382. Doi: 10.1182/blood-2021-148593

Demanelis K, Sriplung H, Meza R, et al. Differences in childhood leukemia incidence and survival between Southern Thailand and the United States: a population-based analysis. *Pediatr Blood Cancer*. 2015;62(10):1790-1798. doi:10.1002/psc.25571

Den Boer, M. L., van Slegtenhorst, M., De Menezes, R. X., et al., 2009. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *The Lancet. Oncology*, 10(2), 125–134.

Dickerson KM, Qu C, Gao Q, et al. ZNF384 Fusion Oncoproteins Drive Lineage Aberrancy in Acute Leukemia. *Blood Cancer Discov*. 2022;3(3):240-263. doi:10.1158/2643-3230.BCD-21-0163

Dmitriev P, Kiseleva E, Kharchenko O, et al. Dux4 controls migration of mesenchymal stem cells through the Cxcr4-Sdf1 axis. *Oncotarget*. 2016;7(40):65090-65108. doi:10.18632/oncotarget.11368

Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21. doi:10.1093/bioinformatics/bts635

Dobson SM, García-Prat L, Vanner RJ, et al. Relapse-Fated Latent Diagnosis Subclones in Acute B Lineage Leukemia Are Drug Tolerant and Possess Distinct Metabolic Programs. *Cancer Discov*. 2020;10(4):568-587. doi:10.1158/2159-8290.CD-19-1059

Doll R, Wakeford R. Risk of childhood cancer from fetal irradiation. *Br J Radiol*. 1997;70:130-139. doi:10.1259/bjr.70.830.9135438

Ding YY, Stern JW, Jubelirer TF, et al. Clinical efficacy of ruxolitinib and chemotherapy in a child with Philadelphia chromosome-like acute lymphoblastic leukemia with GOLGA5-JAK2 fusion and induction failure. *Haematologica*. 2018;103(9):e427-e431. doi:10.3324/haematol.2018.192088

Dong X, Zhang W, Wu H, et al. Structural basis of DUX4/IGH-driven transactivation. *Leukemia*. 2018;32(6):1466-1476. doi:10.1038/s41375-018-0093-1

Dou H, Chen X, Huang Y, et al. Prognostic significance of P2RY8-CRLF2 and CRLF2 overexpression may vary across risk subgroups of childhood B-cell acute lymphoblastic leukemia. *Genes Chromosomes Cancer*. 2017;56(2):135-146. doi:10.1002/gcc.22421

Enshaei A, O'Connor D, Bartram J, et al. A validated novel continuous prognostic index to deliver stratified medicine in pediatric acute lymphoblastic leukemia [published correction appears in *Blood*. 2020 Sep 17;136(12):1468]. *Blood*. 2020;135(17):1438-1446. doi:10.1182/blood.2019003191

Fajardo-Gutiérrez A, González-Miranda G, Pachuca-Vázquez A, Allende-López A, Fajardo-Yamamoto LM, Rendón-Macías ME. Cancer incidence and mortality in children in the Mexican Social Security Institute (1996-2013). *Salud Publica Mex*. 2016;58(2):162-170. doi:10.21149/spm.v58i2.7785

Flores-Lujano J, Duarte-Rodríguez DA, Jiménez-Hernández E. Persistently high incidence rates of childhood acute leukemias from 2010 to 2017 in Mexico City: A population study from the MIGICCL. *Front Public Health*. 2022; 10. Published 2022 Sep 14. doi: 10.3389/fpubh.2022.918921

Ford AM, Bennett CA, Price CM, Bruin MC, Van Wering ER, Greaves M. Fetal origins of the TEL-AML1 fusion gene in identical twins with leukemia. *Proc Natl Acad Sci U S A*. 1998;95(8):4584-4588. doi:10.1073/pnas.95.8.4584.

Gale KB, Ford AM, Repp R, et al. Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots. *Proc Natl Acad Sci U S A*. 1997;94(25):13950-13954. doi:10.1073/pnas.94.25.13950

Gao Q, Liang WW, Foltz SM, et al. Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Rep*. 2018;23(1):227-238.e3. doi:10.1016/j.celrep.2018.03.050

GBD 2017 Childhood Cancer Collaborators. The global burden of childhood and adolescent cancer in 2017: an analysis of the Global Burden of Disease Study 2017 [published correction appears in *Lancet Oncol*. 2019 Aug 6;:] [published correction appears in *Lancet Oncol*. 2021 Aug;22(8):e347]. *Lancet Oncol*. 2019;20(9):1211-1225. doi:10.1016/S1470-2045(19)30339-0

Gelelete CB, Pereira SH, Azevedo AM, et al. Overweight as a prognostic factor in children with acute lymphoblastic leukemia. *Obesity (Silver Spring)*. 2011;19(9):1908-1911. doi:10.1038/oby.2011.195

Gocho Y, Kiyokawa N, Ichikawa H, et al. A novel recurrent EP300-ZNF384 gene fusion in B-cell precursor acute lymphoblastic leukemia. *Leukemia*. 2015;29(12):2445-2448. doi:10.1038/leu.2015.111

Gonzalez-Montalvo P, Romo H, Vega-Vega L, et al. Collaborative Risk-Adapted Treatment for Pediatric Acute Lymphoblastic Leukemia in Mexico. *Blood*. 2017;130(1):2116. doi.org/10.1182/blood.V130.Suppl_1.2116.2116.

Greaves M. Pre-natal origins of childhood leukemia. *Rev Clin Exp Hematol*. 2003;7(3):233-245.

Greaves M. The causation of childhood leukemia: a paradox of progress?. *Discov Med*. 2006;6(31):24-28.

Greaves M. A causal mechanism for childhood acute lymphoblastic leukaemia [published correction appears in *Nat Rev Cancer*. 2018 May 30;:]. *Nat Rev Cancer*. 2018;18(8):471-484. doi:10.1038/s41568-018-0015-6

Greaves M, Hughes W. Cancer cell transmission via the placenta. *Evol Med Public Health*. 2018;2018(1):106-115. Published 2018 Apr 14. doi:10.1093/emph/eoy011

Greenblatt SM, Man N, Hamard PJ, et al. CARM1 Is Essential for Myeloid Leukemogenesis but Dispensable for Normal Hematopoiesis. *Cancer Cell*. 2019;35(1):156. doi:10.1016/j.ccell.2018.12.008

Gröbner SN, Worst BC, Weischenfeldt J, et al. The landscape of genomic alterations across childhood cancers [published correction appears in *Nature*. 2018 Jul;559(7714):E10]. *Nature*. 2018;555(7696):321-327. doi:10.1038/nature25480

Gu Z, Churchman M, Roberts K, et al. Genomic analyses identify recurrent MEF2D fusions in acute lymphoblastic leukaemia. *Nat Commun*. 2016;7:13331. Published 2016 Nov 8. doi:10.1038/ncomms13331

- Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet.* 2019;51(2):296-307. doi:10.1038/s41588-018-0315-5
- Gupta S, Howard SC, Hunger SP, et al. Treating Childhood Cancer in Low- and Middle-Income Countries. In: Gelband H, Jha P, Sankaranarayanan R, Horton S, eds. *Cancer: Disease Control Priorities, Third Edition (Volume 3)*. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; November 1, 2015.
- Harvey RC, Mullighan CG, Chen IM, et al. Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. *Blood.* 2010;115(26):5312-5321. doi:10.1182/blood-2009-09-245944
- Harvey RC, Mullighan CG, Wang X, et al. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. *Blood.* 2010;116(23):4874-4884. doi:10.1182/blood-2009-08-239681
- Hein D, Dreisig K, Metzler M, et al. The preleukemic TCF3-PBX1 gene fusion can be generated in utero and is present in ≈0.6% of healthy newborns. *Blood.* 2019;134(16):1355-1358. doi:10.1182/blood.2019002215
- Herglotz J, Unrau L, Hauschildt F, et al. Essential control of early B-cell development by Mef2 transcription factors. *Blood.* 2016;127(5):572-581. doi:10.1182/blood-2015-04-643270
- Herold T, Schneider S, Metzler KH, et al. Adults with Philadelphia chromosome-like acute lymphoblastic leukemia frequently have IGH-CRLF2 and JAK2 mutations, persistence of minimal residual disease and poor prognosis. *Haematologica.* 2017;102(1):130-138. doi:10.3324/haematol.2015.136366
- Hirabayashi S, Butler ER, Ohki K, et al. Clinical characteristics and outcomes of B-ALL with ZNF384 rearrangements: a retrospective analysis by the Ponte di Legno Childhood ALL Working Group. *Leukemia.* 2021;35(11):3272-3277. doi:10.1038/s41375-021-01199-0
- Hirabayashi S, Ohki K, Nakabayashi K, et al. ZNF384-related fusion genes define a subgroup of childhood B-cell precursor acute lymphoblastic leukemia with a characteristic immunotype. *Haematologica.* 2017;102(1):118-129. doi:10.3324/haematol.2016.151035
- Hjalgrim LL, Westergaard T, Rostgaard K, et al. Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. *Am J Epidemiol.* 2003;158(8):724-735. doi:10.1093/aje/kwg210
- Hu S, Qian M, Zhang H, et al. Whole-genome noncoding sequence analysis in T-cell acute lymphoblastic leukemia identifies oncogene enhancer mutations. *Blood.* 2017;129(24):3264-3268. doi:10.1182/blood-2017-03-771162
- Huether R, Dong L, Chen X, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat Commun.* 2014;5:3630. Published 2014 Apr 8. doi:10.1038/ncomms4630

Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. *N Engl J Med*. 2015;373(16):1541-1552. doi:10.1056/NEJMra1400972

Hunger SP, Raetz EA. How I treat relapsed acute lymphoblastic leukemia in the pediatric population. *Blood*. 2020;136(16):1803-1812. doi:10.1182/blood.2019004043

Iacobucci I, Li Y, Roberts KG, et al. Truncating Erythropoietin Receptor Rearrangements in Acute Lymphoblastic Leukemia. *Cancer Cell*. 2016;29(2):186-200. doi:10.1016/j.ccell.2015.12.013

IARC Monographs on the Identification of Carcinogenic Hazards to Humans, 2020.

Imamura T, Kiyokawa N, Kato M, et al. Characterization of pediatric Philadelphia-negative B-cell precursor acute lymphoblastic leukemia with kinase fusions in Japan. *Blood Cancer J*. 2016;6(5):e419. Published 2016 May 13. doi:10.1038/bcj.2016.28

Inaba H, Azzato EM, Mullighan CG. Integration of Next-Generation Sequencing to Treat Acute Lymphoblastic Leukemia with Targetable Lesions: The St. Jude Children's Research Hospital Approach. *Front Pediatr*. 2017;5:258. Published 2017 Dec 4. doi:10.3389/fped.2017.00258

Inaba H, Greaves M, Mullighan CG. Acute lymphoblastic leukaemia. *Lancet*. 2013;381(9881):1943-1955. doi:10.1016/S0140-6736(12)62187-4

Inaba H, Pui CH. Advances in the Diagnosis and Treatment of Pediatric Acute Lymphoblastic Leukemia. *J Clin Med*. 2021;10(9):1926. Published 2021 Apr 29. doi:10.3390/jcm10091926

Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 2003;4(2):249-264. doi:10.1093/biostatistics/4.2.249

Jann JC, Streuer A, Hecht A, et al. RNA-sequencing of acute promyelocytic leukemia primary blasts reveals novel molecular biomarkers of early death events. *Leuk Lymphoma*. 2020;61(13):3066-3077. doi:10.1080/10428194.2020.1797006

Jeha S, Choi J, Roberts KG, et al. Clinical significance of novel subtypes of acute lymphoblastic leukemia in the context of minimal residual disease-directed therapy. *Blood Cancer Discov*. 2021;2(4):326-337. doi:10.1158/2643-3230.BCD-20-0229

Jeha S, Pei D, Choi J, et al. Improved CNS Control of Childhood Acute Lymphoblastic Leukemia Without Cranial Irradiation: St Jude Total Therapy Study 16. *J Clin Oncol*. 2019;37(35):3377-3391. doi:10.1200/JCO.19.01692

Jiménez-Hernández E, Jaimes-Reyes EZ, Arellano-Galindo J, et al. Survival of Mexican Children with Acute Lymphoblastic Leukaemia under Treatment with the Protocol from the Dana-Farber Cancer Institute 00-01. *Biomed Res Int* 2015;2015:576950.

Jiménez-Hernández E, Fajardo-Gutiérrez A, Núñez-Enriquez JC, et al. A greater birthweight increases the risk of acute leukemias in Mexican children-experience from the Mexican Interinstitutional Group for the Identification of the Causes of Childhood Leukemia (MIGICCL). *Cancer Med*. 2018;7(4):1528-1536. doi:10.1002/cam4.1414

- Jiménez-Morales S, Miranda-Peralta E, Saldaña-Alvarez Y, et al. BCR-ABL, ETV6-RUNX1 and E2A-PBX1: prevalence of the most common acute lymphoblastic leukemia fusion genes in Mexican patients. *Leuk Res.* 2008;32(10):1518-1522. doi:10.1016/j.leukres.2008.03.021
- Juárez-Avendaño G, Luna-Silva NC, Chargoy-Vivaldo E, et al. Poor Prognosis Biomolecular Factors Are Highly Frequent in Childhood Acute Leukemias From Oaxaca, Mexico. *Technol Cancer Res Treat.* 2020; 19:1533033820928436. doi:10.1177/1533033820928436
- Juárez-Velázquez MDR, Moreno-Lorenzana DL, Martínez Anaya DA, et al. High occurrence of CRLF2 abnormalities in Mexican children with B-cell acute lymphoblastic leukemia. *Cytokine.* 2022; 155:155896. doi:10.1016/j.cyto.2022.155896
- Jung M, Schieck M, Hofmann W, et al. Frequency and prognostic impact of PAX5 p.P80R in pediatric acute lymphoblastic leukemia patients treated on an AIEOP-BFM acute lymphoblastic leukemia protocol. *Genes Chromosomes Cancer.* 2020;59(11):667-671. doi:10.1002/gcc.22882
- Jurado S, Fedl AS, Jaritz M, et al. The PAX5-JAK2 translocation acts as dual-hit mutation that promotes aggressive B-cell leukemia via nuclear STAT5 activation. *EMBO J.* 2022;41(7):e108397. doi:10.15252/emboj.2021108397
- Kager L, Lion T, Attarbaschi A, et al. Incidence and outcome of TCF3-PBX1-positive acute lymphoblastic leukemia in Austrian children. *Haematologica.* 2007;92(11):1561-1564. doi:10.3324/haematol.11239
- Kahn JM, Cole PD, Blonquist TM, et al. An investigation of toxicities and survival in Hispanic children and adolescents with ALL: Results from the Dana-Farber Cancer Institute ALL Consortium protocol 05-001. *Pediatr Blood Cancer.* 2018;65(3):10.1002/psc.26871. doi:10.1002/psc.26871
- Kangaspeska S, Hultsch S, Edgren H, Nicorici D, Murumägi A, Kallioniemi O. Reanalysis of RNA-sequencing data reveals several additional fusion genes with multiple isoforms. *PLoS One.* 2012;7(10):e48745. doi:10.1371/journal.pone.0048745
- Kong SG, Seo JH, Jun SE, Lee BK, Lim YT. Childhood acute lymphoblastic leukemia with hyperleukocytosis at presentation. *Blood Res.* 2014;49(1):29-35. doi:10.5045/br.2014.49.1.29
- Kruse A, Abdel-Azim N, Kim HN, et al. Minimal Residual Disease Detection in Acute Lymphoblastic Leukemia. *Int J Mol Sci.* 2020;21(3):1054. Published 2020 Feb 5. doi:10.3390/ijms21031054
- Kumar S, Vo AD, Qin F, Li H. Comparative assessment of methods for the fusion transcripts detection from RNA-Seq data. *Sci Rep.* 2016;6:21597. Published 2016 Feb 10. doi:10.1038/srep21597
- La Starza R, Aventin A, Crescenzi B, et al. CIZ gene rearrangements in acute leukemia: report of a diagnostic FISH assay and clinical features of nine patients. *Leukemia.* 2005;19(9):1696-1699. doi:10.1038/sj.leu.2403842
- Larson RA. Chapter 91 Acute Lymphoblastic Leukemia, En: Kausahnsky, K., & Williams W.J., Williams Hematology. 2016. McGraw-Hill Medical, New York, EEUU. Novena Edición, 1505-1526

- Larsson SC, Wolk A. Overweight and obesity and incidence of leukemia: a meta-analysis of cohort studies. *Int J Cancer*. 2008;122(6):1418-1421. doi:10.1002/ijc.23176
- Lausten-Thomsen U, Madsen HO, Vestergaard TR, Hjalgrim H, Nersting J, Schmiegelow K. Prevalence of t(12;21)[ETV6-RUNX1]-positive cells in healthy neonates. *Blood*. 2011;117(1):186-189. doi:10.1182/blood-2010-05-282764
- Lee SHR, Antillon-Klussmann F, Pei D, et al. Association of Genetic Ancestry With the Molecular Subtypes and Prognosis of Childhood Acute Lymphoblastic Leukemia [published correction appears in *JAMA Oncol*. 2022 Mar 1;8(3):484]. *JAMA Oncol*. 2022;8(3):354-363. doi:10.1001/jamaoncol.2021.6826
- Li JF, Dai YT, Lilljebjörn H, et al. Transcriptional landscape of B cell precursor acute lymphoblastic leukemia based on an international study of 1,223 cases. *Proc Natl Acad Sci U S A*. 2018;115(50):E11711-E11720. doi:10.1073/pnas.1814397115
- Li Z, Lee SHR, Chin WHN, et al. Distinct clinical characteristics of DUX4- and PAX5-altered childhood B-lymphoblastic leukemia. *Blood Adv*. 2021;5(23):5226-5238. doi:10.1182/bloodadvances.2021004895
- Li Y, Wu B, Liu H, et al. Structural basis for multiple gene regulation by human DUX4. *Biochem Biophys Res Commun*. 2018;505(4):1161-1167. doi:10.1016/j.bbrc.2018.10.056
- Lilljebjörn H, Fioretos T. New oncogenic subtypes in pediatric B-cell precursor acute lymphoblastic leukemia. *Blood*. 2017;130(12):1395-1401. doi:10.1182/blood-2017-05-742643
- Lilljebjörn H, Henningsson R, Hyrenius-Wittsten A, et al. Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. *Nat Commun*. 2016;7:11790. Published 2016 Jun 6. doi:10.1038/ncomms11790
- Linares Ballesteros A, Yunis LK, García J, et al. Philadelphia-like acute lymphoblastic leukemia: Characterization in a pediatric cohort in a referral center in Colombia. *Cancer Rep (Hoboken)*. 2022;5(5):e1587. doi:10.1002/cnr2.1587
- Liu YF, Wang BY, Zhang WN, et al. Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. *EBioMedicine*. 2016;8:173-183. doi:10.1016/j.ebiom.2016.04.038
- Locher IJ, Arindrarto W, Borràs DM, et al. Fusion Transcripts without Corresponding Cytogenetic Abnormalities in Acute Myeloid Leukemia: Implications for AML Pathogenesis. *Blood*. 2017;130(1):2703. doi.org/10.1182/blood.V130.Suppl_1.2703.2703.
- Loh ML, DelRocco N, Borowitz M, et al. Enhanced Risk Stratification of 21,178 Children, Adolescents, and Young Adults with Acute Lymphoblastic Leukemia (ALL) Incorporating White Blood Count (WBC), Age, and Minimal Residual Disease (MRD) at Day 8 and 29 As Continuous Variables: A Children's Oncology Group (COG) Report. *Blood*. 2020;136(1): 39-40. doi: 10.1182/blood-2020-139081
- Lu J, McKinsey TA, Nicol RL, Olson EN. Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases. *Proc Natl Acad Sci U S A*. 2000;97(8):4070-4075. doi:10.1073/pnas.080064097

- Ma Y, Dobbins SE, Sherborne AL, et al. Developmental timing of mutations revealed by whole-genome sequencing of twins with acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*. 2013;110(18):7429-7433. doi:10.1073/pnas.1221099110
- Ma X, Liu Y, Liu Y, et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature*. 2018;555(7696):371-376. doi:10.1038/nature25795
- Machado HE, Mitchell E, Øbro NF, et al. Diverse mutational landscapes in human lymphocytes [published online ahead of print, 2022 Aug 10]. *Nature*. 2022;10.1038/s41586-022-05072-7. doi:10.1038/s41586-022-05072-7
- Magrath I, Steliarova-Foucher E, Epelman S, et al. Paediatric cancer in low-income and middle-income countries. *Lancet Oncol*. 2013;14(3):e104-e116. doi:10.1016/S1470-2045(13)70008-1
- Mansour MR, Abraham BJ, Anders L, et al. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science*. 2014;346(6215):1373-1377. doi:10.1126/science.1259037
- Marshall GM, Carter DR, Cheung BB, et al. The prenatal origins of cancer. *Nat Rev Cancer*. 2014;14(4):277-289. doi:10.1038/nrc3679
- Martín-Trejo JA, Núñez-Enríquez JC, Fajardo-Gutiérrez A, et al. Early mortality in children with acute lymphoblastic leukemia in a developing country: the role of malnutrition at diagnosis. A multicenter cohort MIGICCL study. *Leuk Lymphoma*. 2017;58(4):898-908. doi:10.1080/10428194.2016.1219904
- Martínez-Anaya D, Moreno-Lorenzana D, Reyes-León A, et al. Characterization of Philadelphia-like Pre-B Acute Lymphoblastic Leukemia: Experiences in Mexican Pediatric Patients. *Int J Mol Sci*. 2022;23(17):9587. Published 2022 Aug 24. doi:10.3390/ijms23179587
- Mata-Rocha M, Rangel-López A, Jimenez-Hernandez E, et al. Low Prevalence of ETV6::RUNX1 Fusion Gene in a Hispanic Population. *Front Pediatr*. 2022;10:837656. Published 2022 May 24. doi:10.3389/fped.2022.837656
- Maude SL, Dolai S, Delgado-Martin C, et al. Efficacy of JAK/STAT pathway inhibition in murine xenograft models of early T-cell precursor (ETP) acute lymphoblastic leukemia. *Blood*. 2015;125(11):1759-1767. doi:10.1182/blood-2014-06-580480
- McClure BJ, Heatley SL, Kok CH, et al. Pre-B acute lymphoblastic leukaemia recurrent fusion, EP300-ZNF384, is associated with a distinct gene expression. *Br J Cancer*. 2018;118(7):1000-1004. doi:10.1038/s41416-018-0022-0
- McHale CM, Wiemels JL, Zhang L, et al. Prenatal origin of TEL-AML1-positive acute lymphoblastic leukemia in children born in California. *Genes Chromosomes Cancer*. 2003;37(1):36-43. doi:10.1002/gcc.10199
- Medina-Sansón, A. (2016). Chapter 1 Introduction: Childhood Leukemia. En Mejía-Aranguré JM (editor), *Etiology of Acute Leukemias in Children* (pp. 1-48). Springer International Publishing Switzerland. Disponible en línea en: <http://www.springer.com/la/book/9783319057972>.

Mejía-Aranguré JM. Molecular epidemiology of acute leukemia in children: causal model, interaction of three factors-susceptibility, environmental exposure and vulnerability period. *Bol Med Hosp Infant Mex*. 2016;73(1):55-63. doi:10.1016/j.bmhmx.2015.12.005

Meyer C, Burmeister T, Gröger D, et al. The MLL recombinome of acute leukemias in 2017. *Leukemia*. 2018;32(2):273-284. doi:10.1038/leu.2017.213

Mittal P, Roberts CWM. The SWI/SNF complex in cancer - biology, biomarkers and therapy. *Nat Rev Clin Oncol*. 2020;17(7):435-448. doi:10.1038/s41571-020-0357-3

Morak M, Attarbaschi A, Fischer S, et al. Small sizes and indolent evolutionary dynamics challenge the potential role of P2RY8-CRLF2-harboring clones as main relapse-driving force in childhood ALL [published correction appears in *Blood*. 2013 Aug 15;122(7):1328]. *Blood*. 2012;120(26):5134-5142. doi:10.1182/blood-2012-07-443218

Mori H, Colman SM, Xiao Z, et al. Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc Natl Acad Sci U S A*. 2002;99(12):8242-8247. doi:10.1073/pnas.112218799

Möricke A, Reiter A, Zimmermann M, et al. Risk-adjusted therapy of acute lymphoblastic leukemia can decrease treatment burden and improve survival: treatment results of 2169 unselected pediatric and adolescent patients enrolled in the trial ALL-BFM 95 [published correction appears in *Blood*. 2009 Apr 30;113(18):4478. Dosage error in article text]. *Blood*. 2008;111(9):4477-4489. doi:10.1182/blood-2007-09-112920

Mullighan CG, Collins-Underwood JR, Phillips LA, et al. Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. *Nat Genet*. 2009;41(11):1243-1246. doi:10.1038/ng.469

Mullighan CG, Miller CB, Radtke I, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature*. 2008;453(7191):110-114. doi:10.1038/nature06866

Mullighan CG, Su X, Zhang J, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med*. 2009;360(5):470-480. doi:10.1056/NEJMoa0808253

Muñoz-Aguirre P, Huerta-Gutierrez R, Zamora S, et al. Acute Lymphoblastic Leukaemia Survival in Children Covered by Seguro Popular in Mexico: A National Comprehensive Analysis 2005-2017. *Health Syst Reform*. 2021;7(1):e1914897. doi:10.1080/23288604.2021.1914897

Neff T, Armstrong SA. Recent progress toward epigenetic therapies: the example of mixed lineage leukemia. *Blood*. 2013;121(24):4847-4853. doi:10.1182/blood-2013-02-474833

Nicorici D, Satalan M, Edgren H, et al., 2014. FusionCatcher—a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. 2014. doi:10.1101/011650.

Niswander LM, Loftus JP, Lainey É, et al. Therapeutic potential of ruxolitinib and ponatinib in patients with EPOR-rearranged Philadelphia chromosome-like acute lymphoblastic leukemia. *Haematologica*. 2021;106(10):2763-2767. Published 2021 Oct 1. doi:10.3324/haematol.2021.278697

Novakova M, Zaliova M, Fiser K, et al. DUX4r, ZNF384r and PAX5-P80R mutated B-cell precursor acute lymphoblastic leukemia frequently undergo monocytic switch. *Haematologica*. 2021;106(8):2066-2075. Published 2021 Aug 1. doi:10.3324/haematol.2020.250423

Núñez-Enríquez JC, Gil-Hernández AE, Jiménez-Hernández E, et al. Overweight and obesity as predictors of early mortality in Mexican children with acute lymphoblastic leukemia: a multicenter cohort study. *BMC Cancer*. 2019;19(1):708. Published 2019 Jul 18. doi:10.1186/s12885-019-5878-8

O'Connor D, Enshaei A, Bartram J, et al. Genotype-Specific Minimal Residual Disease Interpretation Improves Stratification in Pediatric Acute Lymphoblastic Leukemia. *J Clin Oncol*. 2018;36(1):34-43. doi:10.1200/JCO.2017.74.0449

Ohki K, Kiyokawa N, Saito Y, et al. Clinical and molecular characteristics of MEF2D fusion-positive B-cell precursor acute lymphoblastic leukemia in childhood, including a novel translocation resulting in MEF2D-HNRNP1 gene fusion. *Haematologica*. 2019;104(1):128-137. doi:10.3324/haematol.2017.186320

Okuyama, K., Strid, T., Kuruvilla, J., et al., 2019. PAX5 is part of a functional transcription factor network targeted in lymphoid leukemia. *PLoS genetics*, 15(8), e1008280. <https://doi.org/10.1371/journal.pgen.1008280>

Omar AA, Basiouny L, Elnoby AS, Zaki A, Abouzid M. St. Jude Total Therapy studies from I to XVII for childhood acute lymphoblastic leukemia: a brief review. *J Egypt Natl Canc Inst*. 2022;34(1):25. Published 2022 Jun 13. doi:10.1186/s43046-022-00126-3

Orgel E, Tucci J, Alhushki W, et al. Obesity is associated with residual leukemia following induction therapy for childhood B-precursor acute lymphoblastic leukemia. *Blood*. 2014;124(26):3932-3938. doi:10.1182/blood-2014-08-595389

Oskarsson T, Söderhäll S, Arvidson J, et al. Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome. *Haematologica*. 2016;101(1):68-76. doi:10.3324/haematol.2015.131680

Perez-Andreu V, Roberts KG, Harvey RC, et al. Inherited GATA3 variants are associated with Ph-like childhood acute lymphoblastic leukemia and risk of relapse. *Nat Genet*. 2013;45(12):1494-1498. doi:10.1038/ng.2803

Pérez-Saldivar ML, Fajardo-Gutiérrez A, Bernáldez-Ríos R, et al. Childhood acute leukemias are frequent in Mexico City: descriptive epidemiology. *BMC Cancer*. 2011;11:355. Published 2011 Aug 17. doi:10.1186/1471-2407-11-355

Pérez-Vera P, Montero-Ruiz O, Frías S, et al. Detection of ETV6 and RUNX1 gene rearrangements using fluorescence in situ hybridization in Mexican patients with acute lymphoblastic leukemia: experience at a single institution. *Cancer Genet Cytogenet*. 2005;162(2):140-145. doi:10.1016/j.cancergencyto.2005.03.018

Pieters R, de Groot-Kruseman H, Van der Velden V, et al. Successful Therapy Reduction and Intensification for Childhood Acute Lymphoblastic Leukemia Based on Minimal Residual Disease

Monitoring: Study ALL10 From the Dutch Childhood Oncology Group. *J Clin Oncol*. 2016;34(22):2591-2601. doi:10.1200/JCO.2015.64.6364

Potter N, Jones L, Blair H, et al. Single-cell analysis identifies CRLF2 rearrangements as both early and late events in Down syndrome and non-Down syndrome acute lymphoblastic leukaemia. *Leukemia*. 2019;33(4):893-904. doi:10.1038/s41375-018-0297-4

Potthoff MJ, Olson EN. MEF2: a central regulator of diverse developmental programs. *Development*. 2007;134(23):4131-4140. doi:10.1242/dev.008367

Preston DL, Kusumi S, Tomonaga M, et al. Cancer incidence in atomic bomb survivors. Part III. Leukemia, lymphoma and multiple myeloma, 1950-1987 [published correction appears in *Radiat Res* 1994 Jul;139(1):129]. *Radiat Res*. 1994;137(2 Suppl):S68-S97.

Prima V, Gore L, Caires A, et al. Cloning and functional characterization of MEF2D/DAZAP1 and DAZAP1/MEF2D fusion proteins created by a variant t(1;19)(q23;p13.3) in acute lymphoblastic leukemia. *Leukemia*. 2005;19(5):806-813. doi:10.1038/sj.leu.2403684

Pui CH, Campana D, Pei D, et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. *N Engl J Med*. 2009;360(26):2730-2741. doi:10.1056/NEJMoa0900386

Pulte D, Gondos A, Brenner H. Trends in survival after diagnosis with hematologic malignancy in adolescence or young adulthood in the United States, 1981-2005. *Cancer*. 2009;115(21):4973-4979. doi:10.1002/cncr.24548

Qiu Y, Wang H, Liao P, et al. Systematic pan-cancer landscape identifies CARM1 as a potential prognostic and immunological biomarker. *BMC Genom Data*. 2022;23(1):7. Published 2022 Jan 16. doi:10.1186/s12863-021-01022-w

Quiroz E, Aldoss I, Pullarkat V, Rego E, Marcucci G, Douer D. The emerging story of acute lymphoblastic leukemia among the Latin American population - biological and clinical implications. *Blood Rev*. 2019;33:98-105. doi:10.1016/j.blre.2018.08.002

Raca G, Abdel-Azim H, Yue F, et al. Increased Incidence of IKZF1 deletions and IGH-CRLF2 translocations in B-ALL of Hispanic/Latino children-a novel health disparity. *Leukemia*. 2021;35(8):2399-2402. doi:10.1038/s41375-021-01133-4

Reche PA, Soumelis V, Gorman DM, et al. Human thymic stromal lymphopoietin preferentially stimulates myeloid cells. *J Immunol*. 2001;167(1):336-343. doi:10.4049/jimmunol.167.1.336

Reshmi SC, Harvey RC, Roberts KG, et al. Targetable kinase gene fusions in high-risk B-ALL: a study from the Children's Oncology Group. *Blood*. 2017;129(25):3352-3361. doi:10.1182/blood-2016-12-758979

Rivera-Luna R, Correa-González C, Altamirano-Alvarez E, et al. Incidence of childhood cancer among Mexican children registered under a public medical insurance program. *Int J Cancer*. 2013;132(7):1646-1650. doi:10.1002/ijc.27771

- Roberts KG. Genetics and prognosis of ALL in children vs adults. *Hematology Am Soc Hematol Educ Program*. 2018;2018(1):137-145. doi:10.1182/asheducation-2018.1.137
- Roberts KG, Li Y, Payne-Turner D, et al. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. *N Engl J Med*. 2014;371(11):1005-1015. doi:10.1056/NEJMoa1403088
- Roberts KG, Morin RD, Zhang J, et al. Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. *Cancer Cell*. 2012;22(2):153-166. doi:10.1016/j.ccr.2012.06.005
- Roberts KG, Gu Z, Payne-Turner D, et al. High Frequency and Poor Outcome of Philadelphia Chromosome-Like Acute Lymphoblastic Leukemia in Adults. *J Clin Oncol*. 2017;35(4):394-401. doi:10.1200/JCO.2016.69.0073
- Roberts KG, Reshmi SC, Harvey RC, et al. Genomic and outcome analyses of Ph-like ALL in NCI standard-risk patients: a report from the Children's Oncology Group. *Blood*. 2018;132(8):815-824. doi:10.1182/blood-2018-04-841676
- Robison LL, Bhatia S. Late-effects among survivors of leukaemia and lymphoma during childhood and adolescence. *Br J Haematol*. 2003;122(3):345-359. doi:10.1046/j.1365-2141.2003.04499.x
- Romo H, Gonzalez-Montalvo P, Arce D, et al. Evaluation of Early Treatment Response Utilizing the MAS-ALL18 Adapted Management Guideline in Four "Mexico in Alliance with St. Jude" (MAS) Member Hospitals. *Blood*. 2021. 138(1); 1210. doi.org/10.1182/blood-2021-154325.
- Rosales-Rodríguez B, Núñez-Enríquez JC, Velázquez-Wong AC, et al. Copy Number Alterations are Associated with the Risk of Very Early Relapse in Pediatric B-lineage Acute Lymphoblastic Leukemia: A Nested Case-control MIGICCL Study. *Arch Med Res*. 2021;52(4):414-422. doi:10.1016/j.arcmed.2020.12.013
- Russell LJ, Capasso M, Vater I, et al. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. *Blood*. 2009;114(13):2688-2698. doi:10.1182/blood-2009-03-208397
- Russell LJ, Enshaei A, Jones L, et al. IGH@ translocations are prevalent in teenagers and young adults with acute lymphoblastic leukemia and are associated with a poor outcome [published correction appears in *J Clin Oncol*. 2014 Nov 10;32(32):3687] [published correction appears in *J Clin Oncol*. 2015 Feb 20;33(6):672]. *J Clin Oncol*. 2014;32(14):1453-1462. doi:10.1200/JCO.2013.51.3242
- Russell LJ, Jones L, Enshaei A, et al. Characterisation of the genomic landscape of CRLF2-rearranged acute lymphoblastic leukemia. *Genes Chromosomes Cancer*. 2017;56(5):363-372. doi:10.1002/gcc.22439
- Sasaki K, Yamauchi T, Semba Y, et al. Genome-wide CRISPR-Cas9 screen identifies rationally designed combination therapies for CRLF2-rearranged Ph-like ALL. *Blood*. 2022;139(5):748-760. doi:10.1182/blood.2021012976
- Scheeren FA, van Lent AU, Nagasawa M, et al. Thymic stromal lymphopoietin induces early human B-cell proliferation and differentiation. *Eur J Immunol*. 2010;40(4):955-965. doi:10.1002/eji.200939419

Schinnerl D, Fortschegger K, Kauer M, et al. The role of the Janus-faced transcription factor PAX5-JAK2 in acute lymphoblastic leukemia. *Blood*. 2015;125(8):1282-1291. doi:10.1182/blood-2014-04-570960

Schinnerl D, Mejstrikova E, Schumich A, et al. CD371 cell surface expression: a unique feature of DUX4-rearranged acute lymphoblastic leukemia. *Haematologica*. 2019;104(8):e352-e355. doi:10.3324/haematol.2018.214353

Schmäh J, Fedders B, Panzer-Grümayer R, et al. Molecular characterization of acute lymphoblastic leukemia with high CRLF2 gene expression in childhood. *Pediatr Blood Cancer*. 2017;64(10):10.1002/pbc.26539. doi:10.1002/pbc.26539

Schroeder MP, Bastian L, Eckert C, et al. Integrated analysis of relapsed B-cell precursor Acute Lymphoblastic Leukemia identifies subtype-specific cytokine and metabolic signatures. *Sci Rep*. 2019;9(1):4188. Published 2019 Mar 12. doi:10.1038/s41598-019-40786-1

Schultz KR, Pullen DJ, Sather HN, et al. Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG). *Blood*. 2007;109(3):926-935. doi:10.1182/blood-2006-01-024729

Semchenkova A, Mikhailova E, Komkov A, et al. Lineage Conversion in Pediatric B-Cell Precursor Acute Leukemia under Blinatumomab Therapy. *Int J Mol Sci*. 2022;23(7):4019. Published 2022 Apr 5. doi:10.3390/ijms23074019

Senator H. Zur Kenntniss der Leukämie und Pseudoleukämie im Kindesalter. *Berliner Klinische Wochenschrift* 1882; 35:533–6.

Shago M, Ablu O, Hitzler J, Weitzman S, Abdelhaleem M. Frequency and outcome of pediatric acute lymphoblastic leukemia with ZNF384 gene rearrangements including a novel translocation resulting in an ARID1B/ZNF384 gene fusion. *Pediatr Blood Cancer*. 2016;63(11):1915-1921. doi:10.1002/pbc.26116

Sheykhasan M, Manoochehri H, Dama P. Use of CAR T-cell for acute lymphoblastic leukemia (ALL) treatment: a review study. *Cancer Gene Ther*. 2022;29(8-9):1080-1096. doi:10.1038/s41417-021-00418-1

Shore P, Sharrocks AD. The MADS-box family of transcription factors. *Eur J Biochem*. 1995;229(1):1-13. doi:10.1111/j.1432-1033.1995.tb20430.x

Shyamsunder P, Shanmugasundaram M, Mayakonda A, et al. Identification of a novel enhancer of CEBPE essential for granulocytic differentiation. *Blood*. 2019;133(23):2507-2517. doi:10.1182/blood.2018886077

Siegel DA, Henley SJ, Li J, Pollack LA, Van Dyne EA, White A. Rates and Trends of Pediatric Acute Lymphoblastic Leukemia - United States, 2001-2014. *MMWR Morb Mortal Wkly Rep*. 2017;66(36):950-954. Published 2017 Sep 15. doi:10.15585/mmwr.mm6636a3

Slamova L, Starkova J, Fronkova E, et al. CD2-positive B-cell precursor acute lymphoblastic leukemia with an early switch to the monocytic lineage. *Leukemia*. 2014;28(3):609-620. doi:10.1038/leu.2013.354

Smith MA, Seibel NL, Altekruse SF, et al. Outcomes for children and adolescents with cancer: challenges for the twenty-first century. *J Clin Oncol*. 2010;28(15):2625-2634. doi:10.1200/JCO.2009.27.0421

Soulier J, Clappier E, Cayuela JM, et al. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). *Blood*. 2005;106(1):274-286. doi:10.1182/blood-2004-10-3900

Steinherz PG, Gaynon PS, Breneman JC, et al. Cytoreduction and prognosis in acute lymphoblastic leukemia--the importance of early marrow response: report from the Childrens Cancer Group. *J Clin Oncol*. 1996;14(2):389-398. doi:10.1200/JCO.1996.14.2.389

Steliarova-Foucher E. How can global incidence estimates support childhood cancer control?. *Lancet Oncol*. 2019;20(4):460-461. doi:10.1016/S1470-2045(19)30039-7

Steliarova-Foucher E, Colombet M, Ries LAG, Hesselting P, Moreno F, Shin HY, Stiller CA, editors (2017). *International Incidence of Childhood Cancer, Volume III (electronic version)*. Lyon, France: International Agency for Research on Cancer. Disponible en: <http://iicc.iarc.fr/results/>, accessed [1 de Junio de 2021].

Stiller CA. Incidence of childhood leukaemia, Fact sheet 4.1. Standardized incidence rate of leukaemia as defined by ICD-10 codes C90–95 in children aged 0 to 14 years. WHO Europe. 2009. Disponible en línea en: https://www.euro.who.int/__data/assets/pdf_file/0003/97437/4.1.pdf

Stiller CA. Global burden of childhood cancer: growing, but controllable. *Lancet Oncol*. 2019;20(9):1184-1185. doi:10.1016/S1470-2045(19)30424-3

Suzuki K, Okuno Y, Kawashima N, et al. MEF2D-BCL9 Fusion Gene Is Associated With High-Risk Acute B-Cell Precursor Lymphoblastic Leukemia in Adolescents. *J Clin Oncol*. 2016;34(28):3451-3459. doi:10.1200/JCO.2016.66.5547

Talpaz M, Shah NP, Kantarjian H, et al. Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias. *N Engl J Med*. 2006;354(24):2531-2541. doi:10.1056/NEJMoa055229

Tanaka Y, Kawazu M, Yasuda T, et al. Transcriptional activities of DUX4 fusions in B-cell acute lymphoblastic leukemia. *Haematologica*. 2018;103(11):e522-e526. doi:10.3324/haematol.2017.183152

Tasian SK, Loh ML, Hunger SP. Philadelphia chromosome-like acute lymphoblastic leukemia. *Blood*. 2017;130(19):2064-2072. doi:10.1182/blood-2017-06-743252

Taub JW, Konrad MA, Ge Y, et al. High frequency of leukemic clones in newborn screening blood samples of children with B-precursor acute lymphoblastic leukemia. *Blood*. 2002;99(8):2992-2996. doi:10.1182/blood.v99.8.2992

Teuffel O, Betts DR, Dettling M, Schaub R, Schäfer BW, Niggli FK. Prenatal origin of separate evolution of leukemia in identical twins. *Leukemia*. 2004;18(10):1624-1629. doi: 10.1038/sj.leu.2403462.

- Tian L, Li Y, Edmonson MN, et al. CICERO: a versatile method for detecting complex and diverse driver fusions using cancer RNA sequencing data. *Genome Biol.* 2020;21(1):126. Published 2020 May 28. doi:10.1186/s13059-020-02043-x
- Tian L, Shao Y, Nance S, et al. Long-read sequencing unveils IGH-DUX4 translocation into the silenced IGH allele in B-cell acute lymphoblastic leukemia. *Nat Commun.* 2019;10(1):2789. Published 2019 Jun 26. doi:10.1038/s41467-019-10637-8
- Trigg ME, Sather HN, Reaman GH, et al. Ten-year survival of children with acute lymphoblastic leukemia: a report from the Children's Oncology Group. *Leuk Lymphoma.* 2008;49(6):1142-1154. doi:10.1080/10428190802074593
- Tsuzuki S, Yasuda T, Kojima S, et al. Targeting MEF2D-fusion Oncogenic Transcriptional Circuitries in B-cell Precursor Acute Lymphoblastic Leukemia. *Blood Cancer Discov.* 2020;1(1):82-95. Published 2020 Jun 10. doi:10.1158/2643-3230.BCD-19-0080
- van Dongen JJ, van der Velden VH, Brüggemann M, Orfao A. Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. *Blood.* 2015;125(26):3996-4009. doi:10.1182/blood-2015-03-580027
- Vesely C, Frech C, Eckert C, et al. Genomic and transcriptional landscape of P2RY8-CRLF2-positive childhood acute lymphoblastic leukemia. *Leukemia.* 2017;31(7):1491-1501. doi:10.1038/leu.2016.365
- Wang L, Zhao Z, Meyer MB, et al. CARM1 methylates chromatin remodeling factor BAF155 to enhance tumor progression and metastasis [published correction appears in *Cancer Cell.* 2016 Jul 11;30(1):179-180]. *Cancer Cell.* 2014;25(1):21-36. doi:10.1016/j.ccr.2013.12.007
- Wang J, Hu Y, Huang H. Acute lymphoblastic leukemia relapse after CD19-targeted chimeric antigen receptor T cell therapy. *J Leukoc Biol.* 2017;102(6):1347-1356. doi:10.1189/jlb.5RU0817-315R
- Ward ZJ, Yeh JM, Bhakta N, Frazier AL, Atun R. Estimating the total incidence of global childhood cancer: a simulation-based analysis. *Lancet Oncol.* 2019;20(4):483-493. doi:10.1016/S1470-2045(18)30909-4
- Wiemels JL, Cazzaniga G, Daniotti M, et al. Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet.* 1999;354(9189):1499-1503. doi:10.1016/s0140-6736(99)09403-9
- Wiemels JL, Ford AM, Van Wering ER, Postma A, Greaves M. Protracted and variable latency of acute lymphoblastic leukemia after TEL-AML1 gene fusion in utero. *Blood.* 1999;94(3):1057-1062.
- Williams RT, den Besten W, Sherr CJ. Cytokine-dependent imatinib resistance in mouse BCR-ABL+, Arf-null lymphoblastic leukemia. *Genes Dev.* 2007;21(18):2283-2287. doi:10.1101/gad.1588607
- Williams CK, Foroni L, Luzzatto L, Saliu I, Levine A, Greaves MF. Childhood leukaemia and lymphoma: African experience supports a role for environmental factors in leukaemogenesis. *Ecancermedalscience.* 2014;8:478. Published 2014 Nov 6. doi:10.3332/ecancer.2014.478
- Wolman IJ. Parallel responses to chemotherapy in identical twin infants with concordant leukemia. *J Pediatr.* 1962; 60:91-95. doi:10.1016/s0022-3476(62)80014-6

World Health Organization. CureAll framework: WHO global initiative for childhood cancer: increasing access, advancing quality, saving lives. 2021. World Health Organization. Disponible en: <https://apps.who.int/iris/handle/10665/347370>

Wu A, Cheng L, Vance GH, et al. PAX5 P80R mutated acute leukemia followed by genetically-related histiocytic proliferation with clonal IGH gene rearrangements. *Am J Clin Pathol*. 2021;156(1): S97. doi:10.1093/ajcp/aqab191.206

Wu W, Huang X, Cheng J, et al. Conservation and evolution in and among SRF- and MEF2-type MADS domains and their binding sites. *Mol Biol Evol*. 2011;28(1):501-511. doi:10.1093/molbev/msq214

Yaguchi A, Ishibashi T, Terada K, et al. EP300-ZNF384 fusion gene product up-regulates GATA3 gene expression and induces hematopoietic stem cell gene expression signature in B-cell precursor acute lymphoblastic leukemia cells. *Int J Hematol*. 2017;106(2):269-281. doi:10.1007/s12185-017-2220-6

Yang JJ, Cheng C, Devidas M, et al. Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. *Nat Genet*. 2011;43(3):237-241. doi:10.1038/ng.763

Yang H, Zhang H, Luan Y, et al. Noncoding genetic variation in GATA3 increases acute lymphoblastic leukemia risk through local and global changes in chromatin conformation. *Nat Genet*. 2022;54(2):170-179. doi:10.1038/s41588-021-00993-x

Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*. 2002;1(2):133-143. doi:10.1016/s1535-6108(02)00032-6

Yoshimoto T, Tanaka M, Homme M, et al. CIC-DUX4 Induces Small Round Cell Sarcomas Distinct from Ewing Sarcoma. *Cancer Res*. 2017;77(11):2927-2937. doi:10.1158/0008-5472.CAN-16-3351

Youn HD, Chatila TA, Liu JO. Integration of calcineurin and MEF2 signals by the coactivator p300 during T-cell apoptosis. *EMBO J*. 2000;19(16):4323-4331. doi:10.1093/emboj/19.16.4323

Youn HD, Grozinger CM, Liu JO. Calcium regulates transcriptional repression of myocyte enhancer factor 2 by histone deacetylase 4. *J Biol Chem*. 2000;275(29):22563-22567. doi:10.1074/jbc.C000304200

Yuki Y, Imoto I, Imaizumi M, et al. Identification of a novel fusion gene in a pre-B acute lymphoblastic leukemia with t(1;19)(q23;p13). *Cancer Sci*. 2004;95(6):503-507. doi:10.1111/j.1349-7006.2004.tb03240.x

Zaliova M, Stuchly J, Winkowska L, et al. Genomic landscape of pediatric B-other acute lymphoblastic leukemia in a consecutive European cohort. *Haematologica*. 2019;104(7):1396-1406. doi:10.3324/haematol.2018.204974

Zaliova M, Winkowska L, Stuchly J, et al. A novel class of ZNF384 aberrations in acute leukemia. *Blood Adv*. 2021;5(21):4393-4397. doi:10.1182/bloodadvances.2021005318

Zerrouki R, Benhassine T, Bensaada M, Lauzon P, Trabzi A. The complex translocation (9;14;14) involving IGH and CEBPE genes suggests a new subgroup in B-lineage acute lymphoblastic leukemia. *Genet Mol Biol.* 2016;39(1):7-13. doi:10.1590/S1415-475738420140368

Zhang J, Ding L, Holmfeldt L, et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature.* 2012;481(7380):157-163. Published 2012 Jan 11. doi:10.1038/nature10725

Zhang J, McCastlain K, Yoshihara H, et al. Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. *Nat Genet.* 2016;48(12):1481-1489. doi:10.1038/ng.3691

Zhang Q, Shi C, Han L, et al. Inhibition of mTORC1/C2 signaling improves anti-leukemia efficacy of JAK/STAT blockade in CRLF2 rearranged and/or JAK driven Philadelphia chromosome-like acute B-cell lymphoblastic leukemia. *Oncotarget.* 2018;9(8):8027-8041. Published 2018 Jan 17. doi:10.18632/oncotarget.24261

Zhang M, Zhang H, Li Z, et al. Functional, structural, and molecular characterizations of the leukemogenic driver MEF2D-HNRNPUL1 fusion. *Blood.* 2022;140(12):1390-1407. doi:10.1182/blood.2022016241

Zhao X, Wang P, Diedrich JD, et al. Epigenetic activation of the FLT3 gene by ZNF384 fusion confers a therapeutic susceptibility in acute lymphoblastic leukemia. *Nat Commun.* 2022;13(1):5401. Published 2022 Sep 14. doi:10.1038/s41467-022-33143-w

Zhong CH, Prima V, Liang X, et al. E2A-ZNF384 and NOL1-E2A fusion created by a cryptic t(12;19)(p13.3; p13.3) in acute leukemia. *Leukemia.* 2008;22(4):723-729. doi:10.1038/sj.leu.2405084

Zhou Y, Zhang S, Li Z, et al. Maternal benzene exposure during pregnancy and risk of childhood acute lymphoblastic leukemia: a meta-analysis of epidemiologic studies. *PLoS One.* 2014;9(10):e110466. Published 2014 Oct 15. doi:10.1371/journal.pone.0110466

Zhuang Y, Che J, Wu M, et al. Altered pathways and targeted therapy in double hit lymphoma. *J Hematol Oncol.* 2022;15(1):26. Published 2022 Mar 18. doi:10.1186/s13045-022-01249-9

Zuna J, Madzo J, Krejci O, et al. ETV6/RUNX1 (TEL/AML1) is a frequent prenatal first hit in childhood leukemia. *Blood.* 2011;117(1):368-371. doi:10.1182/blood-2010-09-309070

ANEXOS

Artículo de requisito para la obtención del grado



Article

Transcriptome Analysis Identifies LINC00152 as a Biomarker of Early Relapse and Mortality in Acute Lymphoblastic Leukemia

Diego Alberto Bárcenas-López ¹, Juan Carlos Núñez-Enríquez ², Alfredo Hidalgo-Miranda ³, Fredy Omar Beltrán-Anaya ⁴, Didier Ismael May-Hau ⁵, Elva Jiménez-Hernández ⁶, Vilma Carolina Bekker-Méndez ⁷, Janet Flores-Lujano ², Aurora Medina-Sansón ⁸, Edna Liliana Tamez-Gómez ⁹, Víctor Hugo López-García ¹⁰, José Ramón Lara-Ramos ¹¹, Nora Nancy Núñez-Villegas ⁶, José Gabriel Peñaloza-González ¹², Luz Victoria Flores-Villegas ¹³, Raquel Amador-Sánchez ¹⁴, Rosa Martha Espinosa-Elizondo ¹⁵, Jorge Alfonso Martín-Trejo ¹⁶, Martha Margarita Velázquez-Aviña ¹², Laura Elizabeth Merino-Pasaye ¹³, María Luisa Pérez-Saldívar ², David Aldebarán Duarte-Rodríguez ², José Refugio Torres-Nava ¹⁷, Beatriz Cortés-Herrera ¹⁵, Karina Anastacia Solís-Labastida ¹⁶, Ana Itamar González-Ávila ¹⁴, Jessica Denisse Santillán-Juárez ¹⁸, Alejandra Jimena García-Velázquez ¹⁸, Haydee Rosas-Vargas ¹⁹, Minerva Mata-Rocha ¹⁹, Omar Alejandro Sepúlveda-Robles ¹⁹, Juan Manuel Mejía-Aranguré ^{2,20,*} and Silvia Jiménez-Morales ^{3,*}

- ¹ Programa de Doctorado, Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; d.a.barcenas@outlook.com
- ² Unidad de Investigación Médica en Epidemiología Clínica, UMAE Hospital de Pediatría “Dr. Silvestre Frenk Freund”, Centro Médico Nacional “Siglo XXI”, Instituto Mexicano del Seguro Social, Mexico City 06720, Mexico; jcarlos_nu@hotmail.com (J.C.N.-E.); janetflores22@yahoo.com.mx (J.F.-L.); maria_luisa_2000_mx@yahoo.com (M.L.P.-S.); turunci@gmail.com (D.A.D.-R.)
- ³ Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica (INMEGEN), Mexico City 14610, Mexico; ahidalgo@inmegen.gob.mx
- ⁴ Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; frebeltran@hotmail.com
- ⁵ Programa de Maestría en Investigación Clínica Experimental en Salud, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; didier_may@outlook.com
- ⁶ Servicio de Hematología Pediátrica, Hospital General “Gaudencio González Garza”, Centro Médico Nacional “La Raza”, IMSS, Mexico City 02990, Mexico; elvajimenez@yahoo.com (E.J.-H.); nanuvi_2401@yahoo.com.mx (N.N.N.-V.)
- ⁷ Unidad de Investigación Médica en Inmunología e Infectología, Hospital de Infectología “Dr. Daniel Méndez Hernández”, Centro Médico Nacional “La Raza”, IMSS, Mexico City 02990, Mexico; bekkermendez@yahoo.com
- ⁸ Servicio de Hemato-Oncología, Hospital Infantil de México Federico Gómez, Secretaría de Salud (SS), Mexico City 06720, Mexico; auroramolina@aol.com.mx
- ⁹ Servicio de Hemato-Oncología Hospital Infantil de Tamaulipas, Secretaría de Salud (SS), Cd. Victoria Tamaulipas 87070, Mexico; lilianatamez@hotmail.com
- ¹⁰ Servicio de Ortopedia Pediátrica, Hospital Infantil de Tamaulipas, Secretaría de Salud (SS), Cd. Victoria Tamaulipas 87070, Mexico; ortopediatra81@gmail.com
- ¹¹ Departamento de Genética, Hospital Infantil de Tamaulipas, Secretaría de Salud (SS), Cd. Victoria Tamaulipas 87070, Mexico; lara_mayor@hotmail.com
- ¹² Servicio de Onco-Pediatría, Hospital Juárez de México, Secretaría de Salud (SS), Mexico City 07760, Mexico; penaloza_6@yahoo.es (J.G.P.-G.); m_mvvelazquez@yahoo.com.mx (M.M.V.-A.)
- ¹³ Servicio de Hematología Pediátrica, Centro Médico Nacional “20 de Noviembre”, Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), Mexico City 03100, Mexico; victoriabanco@yahoo.com.mx (L.V.F.-V.); sketch0712@gmail.com (L.E.M.-P.)
- ¹⁴ Hospital General Regional 1 “Dr. Carlos McGregor Sánchez Navarro”, IMSS, Mexico City 03103, Mexico; dibs_amador@hotmail.com (R.A.-S.); itamarga@hotmail.com (A.I.G.-Á.)

- ¹⁵ Servicio de Hematología Pediátrica, Hospital General de México “Dr. Eduardo Liceaga”, Secretaría de Salud (SS), Mexico City 06720, Mexico; rmespinosa1605@hotmail.com (R.M.E.-E.); beatrizcortes101087@gmail.com (B.C.-H.)
- ¹⁶ Servicio de Hematología Pediátrica UMAE Hospital de Pediatría “Dr. Silvestre Frenk Freund”, Centro Médico Nacional “Siglo XXI”, IMSS, Mexico City 06720, Mexico; jorge.martintr@imss.gob.mx (J.A.M.-T.); kas_anastacia@yahoo.com (K.A.S.-L.)
- ¹⁷ Servicio de Oncología, Hospital Pediátrico de Moctezuma, Secretaria de Salud del D.F., Mexico City 15530, Mexico; torresoncoped@live.com.mx
- ¹⁸ Servicio de Hemato-Oncología Pediátrica, Hospital Regional No. 1 de Octubre, ISSSTE, Mexico City 07300, Mexico; jessydenise22@hotmail.com (J.D.S.-J.); ale.garciavelazquez@gmail.com (A.J.G.-V.)
- ¹⁹ Unidad de Investigación en Genética Humana, UMAE Hospital de Pediatría “Dr. Silvestre Frenk Freund”, Centro Médico Nacional “Siglo XXI”, IMSS, Mexico City 06720, Mexico; hayrov@gmail.com (H.R.-V.); mine_mr@hotmail.com (M.M.-R.); sero__82@hotmail.com (O.A.S.-R.)
- ²⁰ Coordinación de Investigación en Salud, IMSS, Mexico City 06720, Mexico
- * Correspondence: juan.mejiaa@imss.gob.mx or arangurejm@gmail.com (J.M.M.-A.); sjimenez@inmegen.gob.mx (S.J.-M.); Tel.: +52–55–5350–1900 (ext. 1155) (S.J.-M.)

Received: 30 January 2020; Accepted: 12 March 2020; Published: 13 March 2020



Abstract: Evidence showing the role of long non-coding RNAs (lncRNAs) in leukemogenesis have emerged in the last decade. It has been proposed that these genes can be used as diagnosis and/or prognosis biomarkers in childhood acute lymphoblastic leukemia (ALL). To know if lncRNAs are associated with early relapse and early mortality, a microarray-based gene expression analysis in children with B-lineage ALL (B-ALL) was conducted. Cox regression analyses were performed. Hazard ratios (HR) and 95% confidence intervals (95% CI) were calculated. *LINC00152* and *LINC01013* were among the most differentially expressed genes in patients with early relapse and early mortality. For *LINC00152* high expression, the risks of relapse and death were HR: 4.16 (95% CI: 1.46–11.86) and HR: 1.99 (95% CI: 0.66–6.02), respectively; for *LINC01013* low expression, the risks of relapse and death were HR: 3.03 (95% CI: 1.14–8.05) and HR: 6.87 (95% CI: 1.50–31.48), respectively. These results were adjusted by NCI risk criteria and chemotherapy regimen. The lncRNA–mRNA co-expression analysis showed that *LINC00152* potentially regulates genes involved in cell substrate adhesion and peptidyl–tyrosine autophosphorylation biological processes. The results of the present study point out that *LINC00152* could be a potential biomarker of relapse in children with B-ALL.

Keywords: acute lymphoblastic leukemia; long noncoding RNA; *LINC00152*; *LINC01013*; early relapse; microarray expression analysis

1. Introduction

Acute leukemia is the most common pediatric malignancy worldwide, representing from one-third of all childhood cancers registered in economically developed countries to nearly 50% in developing ones [1–3]. B-lineage acute lymphoblastic leukemia (B-ALL) is the most frequent leukemia subtype, which displays one of the highest incidences and mortality rates in the Mexican population in comparison with other ethnic groups [2,4]. On the other hand, relapses occurring during early stages of treatment have been documented as one of the main obstacles to achieve better survival rates in Mexican children despite the use of the same chemotherapy protocols of developed countries, where survival rates are higher than 90% at 5-years of follow-up [4,5]. Therefore, better clinical and biological prognosis stratification before treatment initiation is mandatory.

Long non-coding RNAs (lncRNAs) transcripts are emerging as potential diagnosis and/or prognosis biomarkers in diverse human diseases, including hematological malignancies [6–10]. They are defined as transcripts longer than 200 nucleotides, and these genes represent the largest set of non-coding

RNAs in the human transcriptome. They also regulate gene expression at the transcriptional, post-transcriptional, and epigenetic levels through their interaction with DNA, RNA, and proteins [9]. In this regard, different studies conducted in children with ALL and some performed in animal models reported that the lncRNAs signatures might be useful to identify molecular ALL subtypes and novel prognosis biomarkers of the disease [11–16]. As an example, from a whole-transcriptome deep sequencing analysis of the Jurkat leukemic T-cell line, an overexpression of T-ALL-R-lncR1 associated with this ALL subtype was found [11]. As well, Fernando et al. studying 44 ALL children by microarray observed that *BALR-2* overexpression increases proliferation, reduces apoptosis, and correlates with poor overall survival (OS) and a low response to prednisone treatment [7]. Another study of whole transcriptome sequencing involving 56 ALL cases reported that the expression inhibition of the lncRNAs *RP11-624C23.1*, *RP11-203E8*, and *RP11-446E9* could confer advantages to ALL cells by influencing biological processes such as apoptosis, cell proliferation and migration, and DNA damage response and resistance to genotoxic stress [14]. To our knowledge, a limited number of studies have explored the association between the expression of lncRNAs and the risk of relapse in children with B-ALL [16,17]. One of the most recent works was performed in 25 pediatric preB-ALL patients, which, by RNA sequencing and DNA methylation, found a set of *DUX4*, Ph-like specific near haploid/high hyperdiploid subtypes and relapse-specific lncRNAs [16]. The aim of the present study is to investigate whether lncRNA expression profiles associate with the development of early relapses and early deaths in Mexican children with B-ALL.

2. Materials and Methods

2.1. Collection of Human Samples

As part of the Mexican Interinstitutional Group for the Identification of the Causes of Childhood Leukemia (MIGICCL), we carried out a multicenter cohort study which included patients under 18 years old diagnosed with preB-ALL from years 2014 to 2016 and followed-up for at least 36 months after diagnosis. Eight public hospitals located in Mexico City participated.

Children with Down syndrome, other ALL (e.g., T-cell, mixed lineage), and patients living outside Mexico City during treatment were not eligible. Diagnosis confirmation was performed by pediatric hematologists/oncologists based on the morphology and immunophenotype of leukemic cells. Bone marrow (BM) samples were obtained before treatment initiation, and all patients received their chemotherapy treatment in the same hospital where they were diagnosed with ALL.

All clinical data (sex, age at diagnosis, white blood cell count, immunophenotype, risk classification group, and chemotherapy regimen) were registered from the patients' medical records. The National Cancer Institute (NCI) risk stratification was employed (standard risk = from 1 to 9.99 years of age or white blood cell (WBC) count $< 50 \times 10^9/L$, and high risk ≤ 1 or ≥ 10 years of age and/or WBC $\geq 50 \times 10^9/L$).

An early relapse was considered when $\geq 5\%$ leukemic blasts were noted in a bone marrow sample during the first 36 months after having achieved complete remission (CR). The existence of lymphoblasts on smears of cerebrospinal fluid (mononuclear cell count $\geq 5/mL$) accompanied (or not) with cranial nerve paralysis-determined central nervous system (CNS) relapse following the first CR. Early mortality was defined as the patient's death during the first 24-months.

The National Scientific Research and Ethics Committee of the Mexican Institute of Social Security approved the protocol with the number R-2013-785-068. Written informed consent was obtained from the children's parents and patients ≥ 8 years old gave their assent (when possible) to be enrolled in the present study.

2.2. Total RNA Isolation, Microarray Processing, and Data Normalization

Patients whose BM were $>70\%$ in blast cells at diagnosis were included. After BM aspiration, samples were processed, and leukemic blasts were treated with TRIzol reagent (Invitrogen Life

Technologies, Carlsbad, CA, USA). Total RNA isolation and expression analysis were performed as previously described [18]. Briefly, RNA was purified from BM-derived leukemic blasts, then quantified using a Nanodrop spectrophotometer ND1000 (Thermo Fisher Scientific, Waltham, MA, USA), and RNA quality was verified using Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). Samples with an RNA integrity number (RIN) ≥ 7.0 were considered for microarray analyses. Affymetrix GeneChip Human Transcriptome Array 2.0 (HTA 2.0, Affymetrix Inc., Santa Clara, CA, USA) and 200 ng of total RNA were used according to the manufacturer's recommendations (Affymetrix Inc., Santa Clara, CA, USA). The HTA 2.0 array was designed to interrogate coding and non-coding transcripts, including lncRNAs. The RNA was converted into cDNA and labeled with the WT Plus Reagents Kit (Affymetrix, Santa Clara, CA, USA) mixed with the Poly-A RNA spike-in controls and hybridized on the array according to the manufacturer's instructions (each hybridization was performed in duplicate). Arrays were washed, stained and scanned using a GeneChip Scanner 3000 7G (Affymetrix, Santa Clara, CA, USA), followed by the transformation of fluorescence data into CEL files using the Affymetrix GeneChip Command Console (AGCC) software. Background correction, probe sets signal integration, and quantile normalization were performed through the robust multi-array average (RMA) algorithm, also implemented in Affymetrix Expression Console (ECS) software, and following the protocols established by the manufacturer (Affymetrix Inc., Santa Clara, CA, USA).

2.3. mRNA and lncRNA Transcriptional Profiling Analysis

To identify mRNA and lncRNA profiles and potential lncRNA associated with early relapse and mortality, a supervised analysis was conducted, including both coding and non-coding probes (44,699 and 22,829 transcript clusters, respectively). To identify lncRNAs annotated with Ensembl, words like processed_transcript, lincRNA, antisense, non_coding, sense_intronic, ncRNA_host, sense_overlapping and 3prime_overlapping_ncrna, uncharacterized LOC, LINC-, FLJ, (Antisense: AS and -Host gene: HG) were used. Non-annotated probes from the array were manually annotated by searching them on the UCSC genome browser (<https://genome.ucsc.edu/>) and LNCipedia (<https://lncipedia.org/>) webpages. Probes mapping pseudogenes were discarded.

BM samples at the time of diagnosis were classified for analyses if the patient had presented (1) relapse, (2) death during follow-up, and (3) hyperleukocytosis at the time of diagnosis ($\geq 100,000 \times \text{mm}^3$). To select differentially expressed genes, a threshold fold change (FC) value greater than 1.2 and a false discovery rate (FDR) < 0.05 were used.

2.4. Quantitative Real-Time PCR Validation

Quantitative RT-PCR (q-PCR) was performed to validate the expression of the two deregulated lncRNAs identified in the subgroup of children who developed relapse. *LINC00152* (ENSG00000222041) and *LINC01013* (ENSG00000228495) were validated using predesigned TaqMan Gene Expression human assays (Hs03654336_m1 and Hs00395149_m1, respectively) and Universal Master Mix II (Thermo Fisher Scientific, Waltham, MA, USA). Quality control and RNA concentrations were determined by using the NanoDrop 100 spectrophotometer (Thermo Fisher Scientific). cDNA was synthesized from 250 ng of total RNA for each sample using OdT primers and the High-Capacity cDNA Reverse Transcription Kit. Reactions were performed in a final volume of 10 μL under the following conditions: at 95 °C for 10 min, followed by 45 cycles at 95 °C for 15 s and 60 °C for 1 min in a QuantStudio 3 (All Thermo Fisher Scientific, Waltham, MA, USA). Fold changes in expression were calculated by using the $2^{-\Delta\Delta\text{Ct}}$ method. Data were normalized using *SCARNA5* as a control reference gene.

2.5. lncRNA-mRNA Co-Expression Analysis and Biological Pathways Analysis

To gain insight into the biological relevance of the lncRNAs associated with relapse, a guilt-by-association approach was followed to investigate their relationship with different biological pathways [19]. Normalized gene expression of mRNAs and lncRNAs of the 62 arrays (23 ER-ALL and 39 NR-ALL cases) were used to construct a correlation matrix (Spearman correlation and *p*-value),

including only the differentially expressed lncRNAs and mRNAs ($FC = 1.5$, $p < 0.05$ and $FDR \leq 0.05$) genes. For each of the differentially expressed lncRNAs, mRNAs according to their correlation coefficient and p -value were ranked. A significant correlation between lncRNAs and mRNAs was considered when Spearman correlation was greater than 60% (positive or negative) and the p -value was ≤ 0.0001 . The significantly correlated mRNAs were also evaluated by the Gene Ontology (GO) enrichment pathway analysis to define their impact in signaling cell processes.

2.6. Validation Cohort

The validation cohort comprised a subset of cases with ALL and Pre-B immunophenotype, aged 1 to 18 years, included in the MIGICCL study between 2010 and 2013, and who had an available, good quality sample for genetic analysis. Patients who relapsed early (cases) were retrospectively identified in the database of the MIGICCL, and they were frequency-matched with patients without relapse (controls) according to age at diagnosis (± 18 months), chemotherapy protocol, NCI risk stratification, the same time of follow-up, and presence/absence of gene fusions. Two controls per case were selected when possible.

2.7. Statistical Analysis

To compare demographic, clinical and molecular characteristics between groups, the chi-square test or Fisher exact tests were calculated when appropriate. In the discovery cohort, to classify the low and high gene expression, the cutoff value was determined according to the median. This because gene expression values were not normally distributed according to the Shapiro–Wilk test ($p < 0.05$). Nonetheless, gene expression values of lncRNAs in the validation cohort (measured by qRT–PCR) displayed a normal distribution. Thus, to classify the low and high gene expression, the cutoff value was determined according to the mean. Student’s t -test was used to compare the means of expression values between groups. Furthermore, survival analyses were carried out using the Kaplan–Meier method. Disease-free survival (DFS) and overall survival (OS) were calculated, and stratification by NCI risk classification was performed. Log-rank tests were obtained; a p -value less than 0.05 was considered as statistically significant. Afterwards, Cox regression analyses were performed. Hazard ratios (crude and adjusted) and 95% confidence intervals were obtained in both the discovery and validation cohorts. Adjusting variables were selected considering their well-known clinical prognostic relevance and their association with relapse in the univariate analysis. These variables were age ($p < 0.001$), NCI risk ($p = 0.004$) and chemotherapy regimen ($p = 0.04$, for the St. Jude Total XIIIB) in the discovery cohort. However, considering that the NCI risk classification includes the age of the patients and that we identified a high correlation (48%) between NCI risk classification and age in the discovery cohort, we have only included the NCI risk classification in the final model to avoid a multicollinearity phenomenon. Consequently, the variables considered for adjusting HRs in the discovery cohort were NCI risk classification and the chemotherapy regimen. In the validation cohort, the variable selected for adjusting in the multivariate analysis was gender ($p = 0.03$).

3. Results

3.1. Clinical and Demographic Features of the Studied Population

3.1.1. Discovery Cohort

A total of sixty-two BM samples collected at diagnosis from children with de novo B-ALL. Thirty-four (54.8%) patients were male; the median age of the population was 97 months (range: 14–195 months). Seven (11.3%) patients were *BCR-ABL1*, 12 (19.35%) were *ETV6-RUNX1*, 4 (6.45%) were *TCF3-PBX1* positive, and 39 were negative to these four common fusion genes. Twenty-three (37.1%) patients with relapse and 39 (62.9%) with no-relapse were analyzed. The median of the

follow-up of patients was 36 months (range: 3.6–50.1 months) after diagnosis confirmation. Relapses occurred in a range from 2 to 36 (mean = 13) months after achieving complete remission.

Of relapsed cases, isolated BM relapse occurred in 17 (73.9%), three (13.04%) had isolated central nervous system (CNS) relapse, and three (13.04%) presented combined relapse (two: BM + CNS and one: BM + testicular). Fifteen (65%) patients relapsed and died early. The mean survival time in the group of patients without relapse was 36 (30 to 43 months); meanwhile, in patients who suffered early relapse, it was 23 (4 to 50) months. There were no statistically significant differences between groups with or without relapse regarding variables such as gender, leukocyte count in peripheral blood at diagnosis, BM, blast percentage at diagnosis, and gene rearrangement frequencies. Significant differences were noted for age at diagnosis, NCI risk categories, and death ($p = 0.0001$, 0.004, and 0.0001, respectively; Table 1).

Table 1. Clinical characteristics of the children with B-lineage acute lymphoblastic leukemia from the discovery group.

Clinical Characteristics	ALL Children $n = 62$		p^*
	Early Relapse		
	No $n = 39$	Yes $n = 23$	
Gender	n (%)	n (%)	0.35
Female	19 (48.7)	9 (39.1)	
Male	20 (51.3)	14 (60.9)	
Age group (years)			0.0001
1–9.99	29 (74.4)	8 (34.8)	
≥10	10 (25.6)	15 (65.2)	
Age at diagnosis (months)			0.002 **
Median (min–max)	73 (14–191)	140 (18–208)	
WBC count at diagnosis ($\times 10^9/L$)			0.75
<10	13 (33.3)	5 (21.7)	
10–49.99	12 (30.8)	9 (39.1)	
50–99.99	5 (12.8)	4 (17.4)	
≥100	9 (23.1)	5 (21.7)	
BM blast (%) at diagnosis			0.75
<90	7 (17.9)	5 (21.7)	
≥90	32 (82.1)	18 (78.3)	
Gene rearrangement			0.26
<i>ETV6-RUNX1</i>	7 (17.9)	5 (21.7)	
<i>TCF3-PBX1</i>	4 (10.3)	0 (0)	
<i>BCR-ABL1</i>	5 (12.8)	2 (10)	
Non-detected	23 (59.0)	16 (68.3)	
NCI risk classification			0.004
Standard	20 (51.3)	4 (14.3)	
High	19 (48.7)	24 (85.7)	
Relapse site			
Isolated BM	-	17 (74)	
Isolated CNS	-	3 (13)	-
BM and CNS	-	2 (8.7)	
BM and testicular	-	1 (4.3)	

Table 1. Cont.

Clinical Characteristics	ALL Children <i>n</i> = 62		<i>p</i> *
	Early Relapse		
	No <i>n</i> = 39	Yes <i>n</i> = 23	
Chemotherapy regimen			
DFCI (reference)	10 (25.6)	2 (8.7)	-
BFM-95	4 (10.3)	2 (8.7)	0.43
St Jude Total XV	16 (41.0)	8 (34.8)	0.3
St Jude Total XIIB	9 (23.1)	11 (47.8)	0.04
Death			
Yes	0 (0)	15 (65.2)	<0.0001
No	39 (100)	8 (34.8)	

WBC: whole blood count; BM: bone marrow; NCI: National Cancer Institute, NIH, USA; CNS: central nervous system; DFCI: Dana Farber Cancer Institute 00-01; BFM-95: Berlin-Frankfurt-Munster-95; * chi square or Fisher exact test when appropriate. ** Mann-Whitney U-test.

3.1.2. Validation Cohort

In the validation cohort, a total of 69 B-ALL cases diagnosed between 2010 to 2013 were included. The median follow-up time was 34.5 months (range: 9.9–84.7 months). Of these, 20 (25.75%) and 49 (74.25%) patients corresponded to patients with early relapse and patients without relapse, respectively. In contrast with the discovery cohort, statistically significant differences were observed regarding gender and peripheral WBC count at diagnosis. Clinical and demographic characteristics of the validation cohort are displayed in Table S1.

3.2. *LINC00152* and *LINC01013* Were Abnormally Regulated in Patients with Early Relapse and Early Death

To identify differentially expressed lncRNAs in patients with poor outcomes, we considered relapse and death as independent events. We detected 57 lncRNAs differentially expressed ($p < 0.05$) between early relapse vs. no relapse groups, of which 27 were up-regulated (FC = 1.22 to 8.47) and the remainder 30 down-regulated (FC = -1.24 to -5.88). *LINC00152* and *LINC01013* (FC = 2.24, $p = 0.0005$ and FC = -5.88, $p = 0.0004$, respectively) were among the most up- and down-regulated; even so, none of the lncRNA maintained the statistical significance after multiple comparison tests (Table S2A).

Regarding lncRNAs associated with early death, a total of 124 lncRNA were abnormally expressed in cases who died, of which 28 were up-regulated and 96 down-regulated (FC 1.2, $p < 0.05$). To note, the *LINC00152* (FC = 2.03, $p = 0.0048$) and *LINC01013* (FC = -7.79, $p = 0.0004$) were among the most up-regulated and down-regulated genes, respectively (Table S2B).

3.3. Up-Regulated lncRNAs in Patients with hyperleukocytosis

Hyperleukocytosis ($\geq 100 \times 10^9/L$ WBC) is considered one of the most important independent risk predictors of relapse in children with ALL. The analysis conducted to identify lncRNAs associated with hyperleukocytosis exhibited a total of 434 lncRNAs differentially expressed among patients with hyperleukocytosis and patients with a WBC less than $100 \times 10^9/L$. Overall, 130 were up-regulated and 304 down-regulated. Some of the lncRNAs that showed the highest expression were *TCONS_00004231* (FC = 13.49, $p = 0.001$), *SOCS2-AS1* (FC = 5.14, $p = 0.01$), and *LOC101927497* (FC = 3.39, $p = 0.0049$) (Table S2C).

3.4. Co-Expression of lncRNA-Coding Genes and Biological Pathways Analysis

To select a set of lncRNAs and coding RNAs that define the profile of patients with early relapse, a supervised-clustering analysis of gene expression using R software (Bioconductor package) was

carried out. A total of 111 (27 lncRNAs and 84 coding genes) differentially expressed genes (FC = 1.5, $p < 0.01$) among groups with and without relapse were identified (Figure 1).

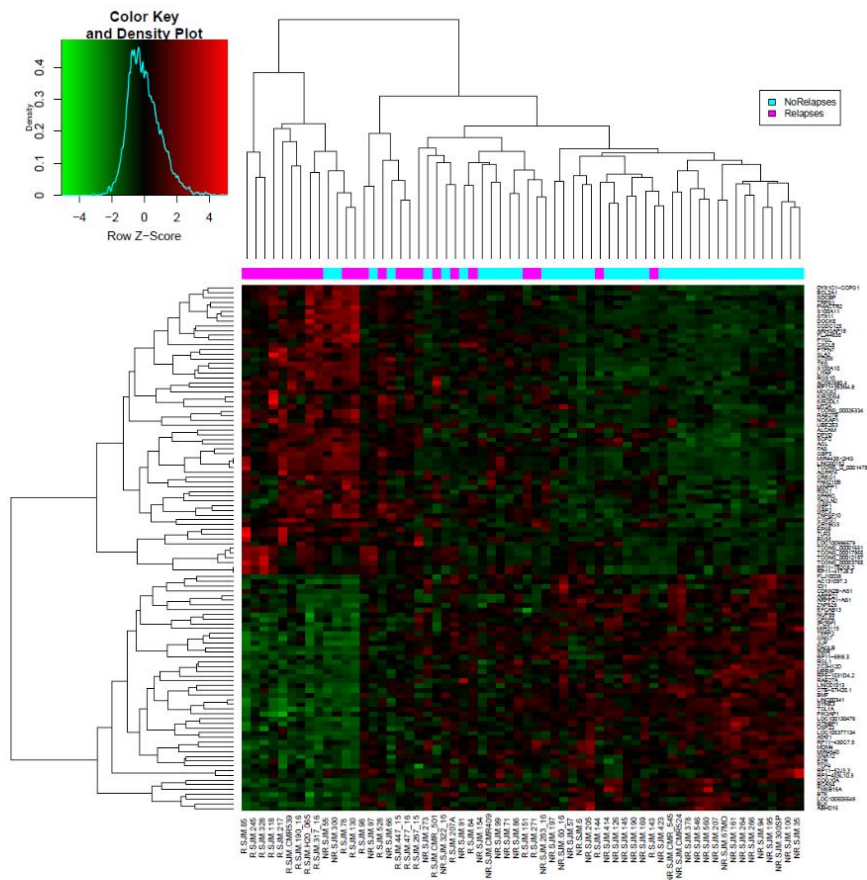


Figure 1. Differentially expressed genes in pediatric patients with acute lymphoblastic leukemia with early relapse vs. non-relapsed cases. Rows display the genes identified, while the columns represent samples (early relapse: pink color and no-relapse: blue). Genes significantly down-regulated are displayed in green and up-regulated genes in red.

In order to detect potential targets of the lncRNAs differentially expressed in groups, we performed a correlation analysis that included all genes. Pairs showing a positive correlation in their expression patterns are displayed in Figure 2.

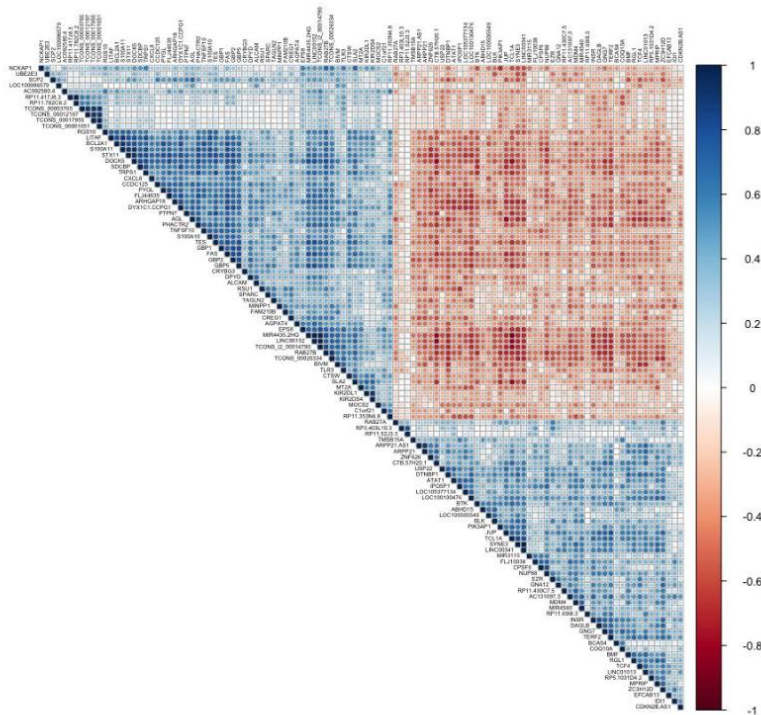


Figure 2. LncRNA–mRNA co-expression. Showing a positive (blue) or negative (brown) correlation between differentially expressed lncRNA and mRNA.

The *TCONS_12_00014790*, *MIR4435.2HG*, *LINC00152*, *TCONS_00026334*, *DYX1C1.CCPG1*, and *LINC01013* were the lncRNAs that showed the highest correlation (Rho 0.6), with most of the abnormally regulated coding genes ($n = 55, 54, 52, 51, 46$, and 32 genes, respectively) (Table S3). Based on the hypothesis that ncRNA and coding genes positively co-expressed are involved in the same biological pathway, we carried out the GO analysis to explore the potential regulatory mechanism of these genes. Our analysis showed that the co-expression correlation network of altered lncRNA–mRNA in relapse was enriched mainly in substrate cellular adhesion, peptidyl-tyrosine autophosphorylation, peptidyl-tyrosine phosphorylation, and regulation of cell proliferation biological processes. *LINC01013* did not show statistical significance with any biological pathway (Figure S1).

3.5. Clinical Association and Survival Analysis

3.5.1. Discovery Cohort

Since *LINC00152* and *LINC01013* have been reported as relevant in the hematopoiesis processes and that both genes exhibited co-expression with coding genes that have been reported as cancer-related, the expression of both genes was used to evaluate their clinical significance. The association among the *LINC00152* and *LINC01013* expression (high or low, according to the median of their expression level) with early relapse and death revealed significant differences in hazard ratios (HRs). In the case of *LINC00152*, its high expression was associated with a higher risk of early relapse (HR = 5.319, 95% CI = 2.48–13.13, log-rank test $p < 0.0001$) and early death (HR = 2.995, 95% CI = 1.146–8.33, log-rank test $p = 0.0193$). Meanwhile, *LINC01013* low expression was also associated with early relapse (HR = 4.019,

95% CI = 1.93–10.15 log-rank test $p < 0.0009$) and death (HR = 9.079, CI = 2.40–17.34 log-rank test $p = 0.0004$) (Figure 3).

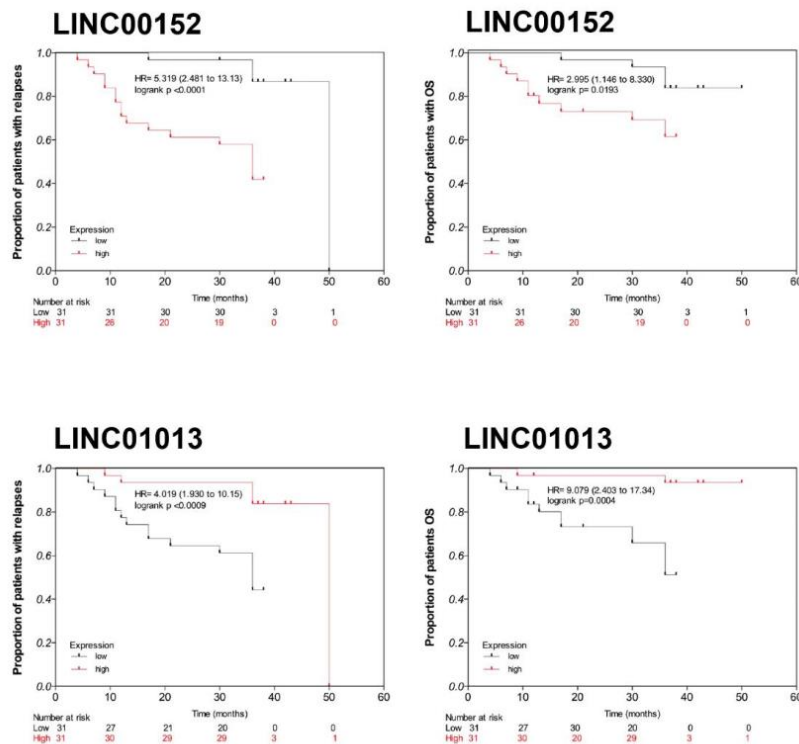


Figure 3. Kaplan–Meier plots (unadjusted) of overall survival for patients with pre-B acute lymphoblastic leukemia. Normalized array expression was used to determine either high (above median) or low (below median) expression of both lncRNAs (p -value was calculated with two-sided log-rank test).

Furthermore, when HRs were adjusted by NCI risk classification and chemotherapy regimen, the associations were the following: (a) for *LINC00152* high expression, the risks of relapse and death were 4.16 (95% CI: 1.46–11.86) and 1.99 (95% CI: 0.66–6.02), respectively, and (b) for *LINC1013* low expression, the risks of relapse and death were HR: 3.03 (95% CI: 1.14–8.05) and HR: 6.87 (95% CI: 1.50–31.48), respectively.

Survival analysis stratified by NCI risk categories showed that in the standard-risk group presenting a low expression level of *LINC00152*, the DFS rate was 93%, while in those having a high expression of *LINC00152*, the disease-free survival (DFS) rate was 62.5%. For those patients classified as high risk according to NCI and having a low expression of *LINC00152*, the DFS was 75%, whereas, in the subgroup presenting high levels of *LINC00152* expression, the DFS was 34.8%. These results were statistically significant (log-rank $p = 0.003$).

When these associations were explored for the *LINC1013*, standard-risk patients presenting high expression levels of this gene showed a DFS rate of 85.7%, whereas in those cases having a low expression of *LINC1013*, the DFS was 77.8% ($p = 0.33$). In the group of patients classified as high risk

and having a high expression of *LINC01013*, the DFS was higher (76.5%) in comparison to children with low expression levels of this gene (31.8%; $p = 0.01$).

It is worth noting that OS rate in patients classified as standard-risk but having a low expression level of *LINC00152* was 93.3%, whereas in those children having a high expression of this gene, the OS was 87.5% ($p = 0.62$). Furthermore, for patients classified as high risk (by NCI criteria) presenting a low expression of *LINC00152*, the OS was 75.0%, and the subgroup with high levels of *LINC00152* expression, the OS was 30.4% ($p = 0.17$).

To note, in the standard-risk group with a high expression level of *LINC01013*, none of the patients died during follow-up, whereas in those having a low expression of this gene, the OS was 77.8% ($p = 0.07$). For patients classified as high risk presenting a high expression of *LINC01013*, the OS was higher compared to the subgroup with low expression levels of *LINC01013* (88.2% and 33.3%, respectively; $p = 0.02$).

3.5.2. Validation Cohort

To validate the results identified in the discovery cohort, the expression of *LINC00152* and *LINC01013* was analyzed by qRT-PCR in an independent cohort (validation cohort) comprised of 69 children with ALL (20 with relapse and 49 without relapse). Given the fact that we used qRT-PCR to evaluate the expression levels of both genes and that the values obtained were normally distributed, low/high expression was categorized according to the mean of expression values (mean: 0.00002218290410 SD = 0.000032891429075 and 0.0363975856084 SD = 0.087986926396332 for *LINC00152* and *LINC01013*, respectively). Remarkably, the expression of both genes exhibited the same tendency as the discovery cohort. The expression levels of *LINC00152* were increased, while they were decreased for *LINC01013* in patients who developed early relapse. Notwithstanding, the statistical significance remained only to *LINC00152* ($p = 0.03$; Figure S2). No associations were identified between levels of expression of *LINC01013* and relapse ($p = 0.16$) or death ($p = 0.43$) in the validation cohort.

In a Kaplan–Meier analysis, a lower DFS rate (50.5%) was observed in the group with a high expression of *LINC00152* in comparison to patients with a low expression (81.8%) (Figure 4). Nonetheless, non-significant differences were observed between levels of expression of *LINC00152* and OS ($p = 0.06$). Similarly, low levels of expression of *LINC01013* were not associated with DFS or OS rates ($p = 0.22$ and $p = 0.59$, respectively).

Furthermore, an increased risk of relapse was observed for the *LINC00152* (cHR: 1.90; 95% CI: 0.75–4.80); however, the confidence intervals were imprecise. When the association was adjusted by gender, the aHR was 1.71 (95% CI: 0.67–4.36). Likewise, no statistically significant association was identified between this gene and death (cHR: 0.98; 95% CI: 0.20–4.87).

By stratifying NCI risk groups in this cohort, none of the patients from the standard-risk group with a low expression level of *LINC00152* relapsed during follow-up, while in those having a high expression of *LINC00152* the DFS rate was 58.4%. Furthermore, for patients classified as high risk according to NCI criteria who presented a low expression of *LINC00152*, the DFS was 74.9%, whereas in the subgroup with high levels of expression for this gene, the DFS was 45.6%. Our results suggest that a high expression of *LINC00152* is associated with lower DFS rates regardless of the NCI risk category ($p = 0.007$). This finding was more evident in the standard-risk group. No associations were observed between a low expression of *LINC01013* and DFS ($p = 0.30$) across NCI groups.

Otherwise, OS analysis showed that in the standard-risk group with a low expression level of *LINC00152*, none of the patients presented early death during follow-up, while in those having a high expression of *LINC00152*, the OS frequency was 93.4%. For patients classified as high risk presenting a low expression of *LINC00152*, the DFS was 89.7%; meanwhile, the DFS was 66.1% in the subgroup with high expression levels for this gene. Thus, our findings reveal that a high expression of *LINC00152* is associated with lower OS rates among NCI categories ($p = 0.03$), being more evident in the high-risk group. No associations were observed between a low expression of *LINC01013* and OS ($p = 0.28$) rates across NCI groups.

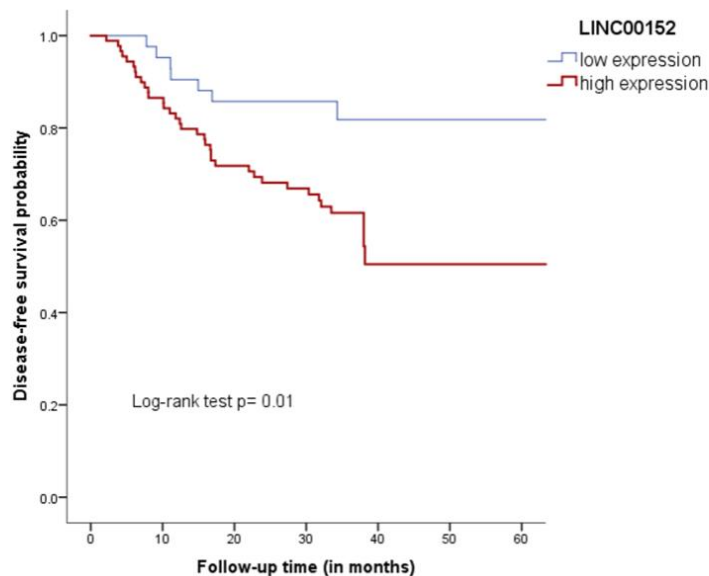


Figure 4. Survival analysis between a low/high expression of LINC00152 with relapse.

4. Discussion

So far, it is predicted that more than 90% of the human genome is actively transcribed, of which only 2% of this genome is coded into proteins and the remaining is transcribed into non-coding genes, mainly lncRNAs [16]. Over the last few years, by using high-throughput technologies for measuring the lncRNAs expression in different cell types (either with normal or pathological conditions), multiple studies have been conducted. Although the molecular mechanisms of most lncRNAs are unknown, increasing evidence suggests that these genes are important regulators of the expression of coding-genes [9]. Among the molecular mechanisms described for these genes are modifications to the spatial conformation of chromosomes, chromatin and DNA modifications, RNA transcription, pre-mRNA splicing, mRNA degradation, mRNA translation, or by competing for microRNA binding [9,20]. Thus, changes in lncRNA expression may contribute to the development and progression of human diseases such as ALL, the most common type of childhood cancer around the globe [1,2,6,9,10]. Despite the advances achieved in the treatment of this disease, relapses still occur in approximately 20% of patients, and it is one of the main obstacles to gain better survival rates in developed populations. Relapse is even greater in Mexican children with ALL (17% during the early stages of treatment) [4].

4.1. LINC00152 and LINC01013 Expression in Acute Lymphoblastic Leukemia

Comparing BM at diagnosis from early relapse vs. no relapse cases, we identified 57 lncRNA differentially expressed, some of which had been previously reported in human tumors (*LOC100130476*, *LINC00152*, *LINC01013*, *MIR4435-2HG*) [21–27]. For instance, it has been referred that down-regulation of *LOC100130476* inhibits proliferation and invasiveness of cancer cells in esophageal squamous cell carcinoma and gastric cardia adenocarcinoma [23,24]. *LINC00152* was abnormally regulated in gastric, pancreatic, hepatic colon, and breast cancer tumors [21,22,26–29]; meanwhile, the *LINC01013* was found to be involved in anaplastic large-cell lymphoma [29]. In our study, the *LINC00152* was one of

the most up-regulated lncRNAs in pre-B ALL cases suffering an early relapse. Additionally, the high expression of *LINC00152* correlated with a high risk for relapse and death.

LINC00152, also known as *CYTOR* (long non-coding RNA cytoskeleton regulator RNA), was recently identified as a cancer-related lncRNA that plays oncogenic roles in several types of cancer [27,30]. A recent meta-analysis exploring the clinical significance and prognosis role of this gene found that an elevated expression of *LINC00152* is associated with metastasis and poor survival in human cancer [27,30]. Additionally, it has been reported that *LINC00152* confers resistance to oxaliplatin in colon cancer [31]. The role of *LINC00152* in biological and pathological processes are continuously being identified, and its location in the nucleus and cytoplasm suggests their involvement in transcription, signaling, and intracellular trafficking [32,33]. Experimental evidence shows that it also regulates cancer-related genes through diverse mechanisms, including epigenetic modifications, miRNAs, and protein interactions [26,33,34]. Data of the present work show that *LINC00152* co-expressed genes are involved in substrate cellular adhesion and peptidyl-tyrosine autophosphorylation biological processes. In vitro and in vivo studies exposed that *LINC00152* promotes tumor progression and epithelial-to-mesenchymal transition process in colon cancer and glioblastoma tumors, respectively, through modulating the *LINC00152*/miR-193a-3p/ERBB4/AKT signaling pathway and inhibiting miR-107 and miR-16 expression [31,34,35]. In addition, in glioblastoma, *LINC00152* overexpression correlated with poor prognosis and an increased invasion through a 3'-hairpin structure [36]. We found that *LINC00152* expression correlates with the *MIR4435-2HG* expression ($R_{h0} = 1$). Both genes were differentially expressed between early relapse and no relapse cases ($FC = 2.22$, $p = 0.0007$ vs. $FC = 2.24$, $p = 0.0005$, respectively). To note, a sequencing analysis of *LINC00152* revealed a high similarity with *MIR4435-2HG*, its paralog gene partner [33]. *MIR4435-2HG* deregulation has been associated with poor prognosis in breast cancer, tumor progression in gastric cancer tissues, and with histological grades and lymph node metastasis in lung cancer [31,37,38]. Functional analysis demonstrated that *MIR4435-2HG* action in cancer could be by its interaction with β -catenin, which promotes cell proliferation and invasion [38]; however, no data regarding *MIR4435-2HG* and ALL have been reported.

To the best of our knowledge, the role of *LINC01013* in childhood acute leukemia had not been explored; nevertheless, data about the relevance of this gene in hematological malignancies had been given in anaplastic large cell lymphoma (ALCL). *LINC01013* was over-expressed in ALCL in comparison with adjacent noncancerous, and in vitro studies demonstrated that its up-regulation enhanced tumor cell invasion through the induction of snail-fibronectin activation cascade [29]. *LINC01013* was downregulated in our cases with early relapse, in contrast with ALCL findings. The discrepancies between both studies could be explained by a dual (oncogene and gene suppressor) role of *LINC01013* in cancer. Dual roles have been observed in other lncRNAs; for instance, *HOXA11-AS*, which is up-regulated (e.g., breast cancer, osteosarcoma, glioma, hepatocarcinoma), but down-regulated (glioblastoma, colorectal cancer, and ovarian cancer) in several types of cancer [39].

Our co-expression analysis revealed that *LINC00152* regulates genes involved in substrate cellular adhesion processes, one of the main biological mechanisms associated with relapse and chemoresistance in ALL. Chemoresistance has been addressed in several studies as one of the most important mechanisms of relapse and death in ALL [40–42]. Taking together, these data and our findings regarding the association between *LINC00152* high expression and the high risk of relapse suggest that *LINC00152* could be potentially involved in chemoresistance; however, functional studies are required to prove these remarks.

4.2. LncRNAs in Hyperleukocytosis

In our study, *SOCS2-AS1* and *LOC101927497* genes were differentially expressed in patients who died and presented hyperleukocytosis in peripheral blood at the time of diagnosis confirmation. Data from prostate cancer suggest that *SOCS2-AS1* lncRNA is involved in the disease by promoting androgen signaling due to the modulation of the epigenetic control for androgen receptors target genes [43]. Regarding *LOC101927497* and its functional role in malignantly transformed human gastric

epithelial cells, it was found that this ncRNA can suppress the proliferation and migration of human gastric epithelial T-cells in vitro by interacting with miR-574–5p. Thus, it was suggested that this gene is an lncRNA suppressor gene in human gastric epithelial cells [44]. To date, no studies have been conducted to identify lncRNAs in patients with hyperleukocytosis in any type of leukemia.

4.3. *LINC00152 and LINC01013 as Potential Biomarkers of Poor Outcomes in Acute Lymphoblastic Leukemia*

The microarray expression analysis showed that *LINC00152* and *LINC01013* were upregulated and downregulated in ER-ALL cases, respectively. Using qRT-PCR, we validated these findings in a subset of patients from the discovery cohort (data not shown).

In addition, we sought to validate microarrays results by qRT-PCR in an independent population (validation cohort). Statistically significant differences in the association between the expression levels of *LINC00152* but not to *LINC01013* and relapse were observed. The purpose of having made a frequency-match instead of analyzing available data in the validation cohort was to control the effect of variables that are currently used in risk stratification and treatment assignment of children with B-ALL; for instance, age of patients, NCI risk classification, and gene rearrangements. Furthermore, given the different protocols used in the hospitals, we have decided to match by chemotherapy regimen. As displayed in Table S1, non-significant differences were observed in these matching variables between relapse and no-relapse groups. The only variable that showed significant differences among the studied groups was the gender of patients. Then, this variable was selected for adjusting the HRs in the validation cohort. Even though the sample size of the discovery and the validation cohorts were similar, there is still clinical and molecular heterogeneity between them, as well as the chemotherapeutic protocols used by the participants' institutions, support therapy, and nutritional status that can influence our results [45]. Therefore, statistical differences between discovery and validation cohorts are probably due to biological reasons.

Other limitations of our work include the small sample size, which, in addition to the high molecular heterogeneity of the ALL, influences the statistical power. Furthermore, we analyzed BM from patients who have >70% (at diagnosis) of leukemia cells, so normal hematopoietic cells could act as a confounding factor. To overcome this problem, blast cells should be sorted. Otherwise, we do not know if the stability of the lncRNAs differs under normal and pathological conditions during the evolution of the disease or their association with treatment response [46]. In this regard, it is necessary to investigate whether these genes have an independent prognostic role regarding the minimal residual disease (MRD) status of ALL patients with relapse; however, MRD is not currently performed in our participating hospitals.

To our knowledge, no previous studies have evaluated the relationship between these genes and the risk of relapse or death in ALL cases. Notwithstanding, *LINC00152* and *LINC01013* expression have been correlated with poor prognosis in diverse human cancer types, including acute myeloid leukemia [27,29,47], supporting the potential role of both genes in ALL as risk predictors of poor outcomes. Thus, in order to gain a better understanding of the role of *LINC00152* and *LINC01013* as biomarkers associated with relapse and death in children with ALL, the sample size should be increased in further investigations.

5. Conclusions

Nowadays, epigenetic factors are considered as one of the most effective mechanisms in the pathogenesis of different malignancies. These factors are reversible elements, which open up the possibility to use them as indicators of the disease, prognosis biomarkers, and important therapy targets. In the present study, the results support the previously reported evidence pointing out that lncRNAs' deregulated expression is linked to ALL and relapse; however, further studies are needed to endorse the value of these genes as independent biomarkers of early relapses and deaths.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/11/3/302/s1>. Figure S1: Biological pathways by guilt-by-association analysis using lncRNA-mRNA co-expression. Figure S2: Quantitative PCR (qPCR) analysis. Relative expression levels of *LINC00152* and *LINC01013* normalized to *SCARNA5*. Table S1: Demographic and clinical characteristics of the validation cohort. Table S2: Differentially expressed lncRNAs among patients with (A) relapse and (B) death and hyperleukocytosis (C). Table S3: Coding genes displaying the higher correlation values between lncRNAs and mRNAs.

Author Contributions: Conceptualization, S.J.-M., and J.M.M.-A.; Methodology, D.A.B.-L., J.C.N.-E., V.C.B.-M., and S.J.-M.; Validation, D.I.M.-H., D.A.B.-L., J.C.N.-E., and V.C.B.-M.; Formal Analysis, D.A.B.-L., D.I.M.-H., J.C.N.-E., and F.O.B.-A.; Investigation, D.A.B.-L., J.C.N.-E., and S.J.-M.; Resources, E.J.-H., V.C.B.-M., J.F.-L., M.L.P.-S., A.M.-S., E.L.T.-G., V.H.L.-G., J.R.L.-R., N.N.N.-V., J.G.P.-G., L.V.F.-V., M.M.V.-A., R.A.-S., A.I.G.-Á., R.M.E.-E., B.C.-H., J.A.M.-T., K.A.S.-L., L.E.M.-P., N.C.R.-Z., J.D.S.-J., A.J.G.-V., H.R.-V., M.M.-R., O.A.S.-R.; S.J.-M., and J.M.M.-A.; Writing—Original Draft Preparation, D.A.B.-L. and S.J.-M.; Writing—Review and Editing, S.J.-M., J.M.M.-A., and A.H.-M.; Supervision, S.J.-M. and J.M.M.-A.; Funding Acquisition, S.J.-M. and J.M.M.-A. Review Final manuscript: All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT); grant numbers Investigación en Fronteras de la Ciencia (IFC)-2016-01-2119, PDCPN2013-01-215726, SALUD 2015-1-262190, CB-2015-1-258042, and FONCICYT/37/2018), the Instituto Mexicano del Seguro Social (grant numbers: FIS/IMSS/PROT/1364, FIS/IMSS/PROT/1548, FIS/IMSS/PROT/895, FIS/IMSS/PROT/1533, FIS/IMSS/PROT/1782, FIS/IMSS/PROT/PRI0/14/031, FIS/IMSS/PROT/PRI0/15/048, FIS/IMSS/PROT/MD15/1504, FIS/IMSS/PROT/G15/1477), and by the National Institute of Genomic Medicine (19/2019/I). The funding body had no role in the design of the study, collection, analysis, and interpretation of the data, or preparation of the manuscript. The Diego Alberto Bárcenas López fellowship was provided by Consejo Nacional de Ciencia y Tecnología CONACyT (CVU 737534).

Acknowledgments: We thank all the patients and their families at hospitals, whose willingness to participate made this study possible. As well, we thank the Microarrays Unit of the National Institute of Genomic Medicine for the excellent technical support provided. This work was submitted in partial fulfillment of the requirements for the PhD degree for Diego Alberto Bárcenas-López at Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México (UNAM).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Barrington-Trimis, J.L.; Cockburn, M.; Metayer, C.; Gauderman, W.J.; Wiemels, J.; McKean-Cowdin, R. Trends in childhood leukemia incidence over two decades from 1992 to 2013. *Int. J. Cancer* **2017**, *140*, 1000–1008. [[CrossRef](#)]
- Rivera-Luna, R.; Velasco-Hidalgo, L.; Zapata-Tarrés, M.; Cárdenas-Cardos, R.; Aguilar-Ortiz, M.R. Current outlook of childhood cancer epidemiology in a middle-income country under a public health insurance program. *Pediatr. Hematol. Oncol.* **2017**, *34*, 43–50. [[CrossRef](#)] [[PubMed](#)]
- Lin, M.S.; Brown, L.M.; Mbulaiteye, S.M.; Check, D.; Ostroumova, E.; Landgren, A.; Devesa, S.S. International long-term trends and recent patterns in the incidence of leukemias and lymphomas among children and adolescents ages 0–19 years. *Int. J. Cancer* **2016**, *138*, 1862–1874. [[CrossRef](#)] [[PubMed](#)]
- Jiménez-Hernández, E.; Jaimes-Reyes, E.Z.; Arellano-Galindo, J.; García-Jiménez, X.; Tiznado-García, H.M.; Dueñas-González, M.T.; Martínez-Villegas, O.; Sánchez-Jara, B.; Bekker-Méndez, V.C.; Ortiz-Torres, M.G.; et al. Survival of Mexican Children with Acute Lymphoblastic Leukaemia under Treatment with the Protocol from the Dana-Farber Cancer Institute 00–01. *Biomed. Res. Int.* **2015**, *2015*, 1–9. [[CrossRef](#)]
- Martín-Trejo, J.A.; Núñez-Enríquez, J.C.; Fajardo-Gutiérrez, A.; Medina-Sansón, A.; Flores-Lujano, J.; Jiménez-Hernández, E.; Amador-Sánchez, R.; Peñaloza-González, J.G.; Álvarez-Rodríguez, F.J.; Bolea-Murga, V.; et al. Early mortality in children with acute lymphoblastic leukemia in a developing country: The role of malnutrition at diagnosis. A multicenter cohort MIGICCL study. *Leuk. Lymphoma* **2017**, *58*, 898–908. [[CrossRef](#)] [[PubMed](#)]
- Fang, K.; Han, B.W.; Chen, Z.H.; Lin, K.Y.; Zeng, C.W.; Li, X.J.; Li, J.H.; Luo, X.Q.; Chen, Y.Q. A distinct set of long non-coding RNAs in childhood MLL-rearranged acute lymphoblastic leukemia: Biology and epigenetic target. *Hum. Mol. Genet.* **2014**, *23*, 3278–3288. [[CrossRef](#)] [[PubMed](#)]
- Fernando, T.R.; Rodríguez-Malave, N.I.; Waters, E.V.; Yan, W.; Casero, D.; Basso, G.; Pigazzi, M.; Rao, D.S. LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia. *Mol. Cancer Res.* **2015**, *13*, 839–851. [[CrossRef](#)] [[PubMed](#)]

8. Cao, L.; Xiao, P.F.; Tao, Y.F.; Hu, S.Y.; Lu, J.; Zhao, W.L.; Li, Z.H.; Wang, N.N.; Wang, J.; Feng, X.; et al. Microarray profiling of bone marrow long non-coding RNA expression in Chinese pediatric acute myeloid leukemia patients. *Oncol. Rep.* **2016**, *35*, 757–770. [[CrossRef](#)]
9. Alvarez-Dominguez, J.R.; Lodish, H.F. Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. *Blood* **2017**, *130*, 1965–1975. [[CrossRef](#)]
10. Lammens, T.; Durinck, K.; Wallaert, A.; Speleman, F.; Van Vlierberghe, P. Long non-coding RNAs in leukemia: Biology and clinical impact. *Curr. Opin. Hematol.* **2017**, *24*, 353–358. [[CrossRef](#)]
11. Zhang, L.; Xu, H.G.; Lu, C. A novel long non-coding RNA T-ALL-R-LncR1 knockdown and Par-4 cooperate to induce cellular apoptosis in T-cell acute lymphoblastic leukemia cells. *Leuk. Lymphoma* **2014**, *55*, 1373–1382. [[CrossRef](#)] [[PubMed](#)]
12. Ghazavi, F.; De Moerloose, B.; Van Looche, W.; Wallaert, A.; Helmsmoortel, H.H.; Ferster, A.; Bakkus, M.; Plat, G.; Delabesse, E.; Uytendaele, A.; et al. Unique long non-coding RNA expression signature in ETV6/RUNX1-driven B-cell precursor acute lymphoblastic leukemia. *Oncotarget* **2016**, *7*, 73769–73780. [[CrossRef](#)] [[PubMed](#)]
13. Wallaert, A.; Durinck, K.; Van Looche, W.; Van de Walle, I.; Matthijssens, F.; Volders, P.J.; Avila-Cobos, F.; Rombaut, D.; Rondou, P.; Mestdagh, P.; et al. Long noncoding RNA signatures define oncogenic subtypes in T-cell acute lymphoblastic leukemia. *Leukemia* **2016**, *30*, 1927–1930. [[CrossRef](#)] [[PubMed](#)]
14. Gioia, R.; Drouin, S.; Ouimet, M.; Caron, M.; St-Onge, P.; Richer, C.; Sinnett, D. LncRNAs downregulated in childhood acute lymphoblastic leukemia modulate apoptosis, cell migration, and DNA damage response. *Oncotarget* **2017**, *8*, 80645–80650. [[CrossRef](#)]
15. Ngoc, P.C.T.; Tan, S.H.; Tan, T.K.; Chan, M.M.; Li, Z.; Yeoh, A.E.J.; Tenen, D.G.; Sanda, T. Identification of novel lncRNAs regulated by the TAL1 complex in T-cell acute lymphoblastic leukemia. *Leukemia* **2018**, *32*, 2138–2151. [[CrossRef](#)]
16. James, A.R.; Schroeder, M.P.; Neumann, M.; Bastian, L.; Eckert, C.; Gökbüget, N.; Sanchez, J.O.; Schlee, C.; Isaakidis, K.; Schwartz, S.; et al. Long non-coding RNAs defining major subtypes of B cell precursor acute lymphoblastic leukemia. *J. Hematol. Oncol.* **2019**, *12*, 8. [[CrossRef](#)]
17. Rosa, R.; Villegas-Ruiz, V.; Caballero-Palacios, M.C.; Pérez-López, E.I.; Murata, C.; Zapata-Tarres, M.; Cárdenas-Cardos, R.; Paredes-Aguilera, R.; Rivera-Luna, R.; Juárez-Méndez, A.S. Expression of ZNF695 Transcript Variants in Childhood B-Cell Acute Lymphoblastic Leukemia. *Genes* **2019**, *10*, 716. [[CrossRef](#)]
18. Núñez-Enríquez, J.C.; Bárcenas-López, D.A.; Hidalgo-Miranda, A.; Jiménez-Hernández, E.; Bekker-Méndez, V.C.; Flores-Lujano, J.; Solís-Labastida, K.A.; Martínez-Morales, G.B.; Sánchez-Muñoz, F.; Espinoza-Hernández, L.E.; et al. Gene Expression Profiling of Acute Lymphoblastic Leukaemia in Children with Very Early Relapse. *Arch. Med. Res.* **2016**, *47*, 644–655. [[CrossRef](#)]
19. Chen, E.Y.; Tan, C.M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G.V.; Clark, N.R.; Ma'ayan, A. Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **2013**, *14*, 128. [[CrossRef](#)]
20. Tehrani, S.S.; Karimian, A.; Parsian, H.; Majidinia, M.; Yousefi, B. Multiple Functions of Long Non-Coding RNAs in Oxidative Stress, DNA Damage Response and Cancer Progression. *J. Cell. Biochem.* **2018**, *119*, 223–236. [[CrossRef](#)]
21. Cao, W.J.; Wu, H.L.; He, B.S.; Zhang, Y.S.; Zhang, Z.Y. Analysis of long non-coding RNA expression profiles in gastric cancer. *World J. Gastroenterol.* **2013**, *19*, 3658–3664. [[CrossRef](#)] [[PubMed](#)]
22. Morlando, M.; Ballarino, M.; Fatica, A. Long Non-Coding RNAs: New Players in Hematopoiesis and Leukaemia. *Front. Med.* **2015**, *2*, 23. [[CrossRef](#)] [[PubMed](#)]
23. Müller, S.; Raulefs, S.; Bruns, P.; Afonso-Grunz, F.; Plötner, A.; Thermann, R.; Jäger, C.; Schlitter, A.M.; Kong, B.; Regel, I.; et al. Next-generation sequencing reveals novel differentially regulated mRNAs, lncRNAs, miRNAs, sdRNAs and a piRNA in pancreatic cancer. *Mol. Cancer* **2015**, *14*, 94. [[CrossRef](#)] [[PubMed](#)]
24. Guo, W.; Dong, Z.; Shi, Y.; Liu, S.; Liang, J.; Guo, Y.; Guo, X.; Shen, S.; Wang, G. Methylation-mediated downregulation of long noncoding RNA LOC100130476 in gastric cardia adenocarcinoma. *Clin. Exp. Metastasis* **2016**, *33*, 497–508. [[CrossRef](#)]
25. Guo, W.; Dong, Z.; Shi, Y.; Liu, S.; Liang, J.; Guo, Y.; Guo, X.; Shen, S.; Shan, B. Aberrant methylation-mediated downregulation of long noncoding RNA LOC100130476 correlates with malignant progression of esophageal squamous cell carcinoma. *Dig. Liver Dis.* **2016**, *48*, 961–969. [[CrossRef](#)] [[PubMed](#)]

26. Ke, D.; Li, H.; Zhang, Y.; An, Y.; Fu, H.; Fang, X.; Zheng, X. The combination of circulating long noncoding RNAs AK001058, INHBA-AS1, MIR4435-2HG, and CEBPA-AS1 fragments in plasma serve as diagnostic markers for gastric cancer. *Oncotarget* **2017**, *8*, 21516–21525. [[CrossRef](#)] [[PubMed](#)]
27. Hu, X.L.; Wang, J.; He, W.; Zhao, P.; Wu, W.Q. Down-regulation of lncRNA Linc00152 suppressed cell viability, invasion, migration, and epithelial to mesenchymal transition, and reversed chemo-resistance in breast cancer cells. *Eur. Rev. Med. Pharmacol. Sci.* **2018**, *22*, 3074–3084. [[CrossRef](#)]
28. Ma, P.; Wang, H.; Sun, J.; Liu, H.; Zheng, C.; Zhou, X.; Lu, Z. LINC00152 promotes cell cycle progression in hepatocellular carcinoma via miR-193a/b-3p/CCND1 axis. *Cell Cycle* **2018**, *17*, 974–984. [[CrossRef](#)]
29. Chung, I.H.; Lu, P.H.; Lin, Y.H.; Tsai, M.M.; Lin, Y.W.; Yeh, C.T.; Lin, K.H. The long non-coding RNA LINC01013 enhances invasion of human anaplastic large-cell lymphoma. *Sci. Rep.* **2017**, *7*, 295. [[CrossRef](#)]
30. Bian, Z.; Zhang, J.; Li, M.; Feng, Y.; Yao, S.; Song, M.; Qi, X.; Fei, B.; Yin, Y.; Hua, D.; et al. Long non-coding RNA LINC00152 promotes cell proliferation, metastasis, and confers 5-FU resistance in colorectal cancer by inhibiting miR-139-5p. *Oncogenesis* **2017**, *6*, 395. [[CrossRef](#)]
31. Yue, B.; Cai, D.; Liu, C.; Fang, C.; Yan, D. Linc00152 Functions as a Competing Endogenous RNA to Confer Oxaliplatin Resistance and Holds Prognostic Values in Colon Cancer. *Mol. Ther.* **2016**, *24*, 2064–2077. [[CrossRef](#)] [[PubMed](#)]
32. Soares, R.J.; Maglieri, G.; Gutschner, T.; Diederichs, S.; Lund, A.H.; Nielsen, B.S.; Holmstrøm, K. Evaluation of fluorescence in situ hybridization techniques to study long non-coding RNA expression in cultured cells. *Nucleic Acids Res.* **2018**, *46*, e4. [[CrossRef](#)] [[PubMed](#)]
33. Nötzold, L.; Frank, L.; Gandhi, M.; Polycarpou-Schwarz, M.; Groß, M.; Gunkel, M.; Beil, N.; Erfle, H.; Harder, N.; Rohr, K.; et al. The long non-coding RNA LINC00152 is essential for cell cycle progression through mitosis in HeLa cells. *Sci. Rep.* **2017**, *7*, 2265. [[CrossRef](#)] [[PubMed](#)]
34. Chen, W.M.; Huang, M.D.; Sun, D.P.; Kong, R.; Xu, T.P.; Xia, R.; Zhang, E.B.; Shu, Y.Q. Long intergenic non-coding RNA 00152 promotes tumor cell cycle progression by binding to EZH2 and repressing p15 and p21 in gastric cancer. *Oncotarget* **2016**, *7*, 9773–9787. [[CrossRef](#)]
35. Liu, X.; Yidayitula, Y.; Zhao, H.; Luo, Y.; Ma, X.; Xu, M. LncRNA LINC00152 promoted glioblastoma progression through targeting the miR-107 expression. *Environ. Sci. Pollut. Res. Int.* **2018**, *25*, 17674–17681. [[CrossRef](#)]
36. Reon, B.J.; Takao-Real-Karia, B.; Kiran, M.; Dutta, A. LINC00152 Promotes Invasion through a 3'-Hairpin Structure and Associates with Prognosis in Glioblastoma. *Mol. Cancer Res.* **2018**, *16*, 1470–1482. [[CrossRef](#)]
37. Miao, Y.; Sui, J.; Xu, S.Y.; Liang, G.Y.; Pu, Y.P.; Yin, L.H. Comprehensive analysis of a novel four-lncRNA signature as a prognostic biomarker for human gastric cancer. *Oncotarget* **2017**, *8*, 75007–75024. [[CrossRef](#)]
38. Qian, H.; Chen, L.; Huang, J.; Wang, X.; Ma, S.; Cui, F.; Luo, L.; Ling, L.; Luo, K.; Zheng, G. The lncRNA MIR4435-2HG promotes lung cancer progression by activating β -catenin signalling. *J. Mol. Med.* **2018**, *96*, 753–764. [[CrossRef](#)]
39. Li, T.; Xu, C.; Cai, B.; Zhang, M.; Gao, F.; Gan, J. Expression and clinicopathological significance of the lncRNA HOXA11-AS in colorectal cancer. *Oncol. Lett.* **2016**, *12*, 4155–4160. [[CrossRef](#)]
40. Barwe, S.P.; Quagliano, A.; Gopalakrishnapillai, A. Eviction from the sanctuary: Development of targeted therapy against cell adhesion molecules in acute lymphoblastic leukaemia. *Semin. Oncol.* **2017**, *44*, 101–112. [[CrossRef](#)]
41. Sun, H.; Zhang, Z.; Luo, W.; Liu, J.; Lou, Y.; Xia, S. NET1 Enhances Proliferation and Chemoresistance in Acute Lymphoblastic Leukaemia Cells. *Oncol. Res.* **2019**, *27*, 935–944. [[CrossRef](#)] [[PubMed](#)]
42. Yu, K.; Wang, J.; Lu, T.; Ma, D.; Wei, D.; Guo, Y.; Cheng, B.; Wang, W.; Fang, Q. Overexpression of heme oxygenase-1 in microenvironment mediates vincristine resistance of B-cell acute lymphoblastic leukaemia by promoting vascular endothelial growth factor secretion. *J. Cell. Biochem.* **2019**, *120*, 17791–17810. [[CrossRef](#)] [[PubMed](#)]
43. Misawa, A.; Takayama, K.; Urano, T.; Inoue, S. Androgen-induced Long Noncoding RNA (lncRNA) SOCS2-AS1 Promotes Cell Growth and Inhibits Apoptosis in Prostate Cancer Cells. *J. Biol. Chem.* **2016**, *291*, 17861–17880. [[CrossRef](#)]
44. Luo, Y.; Liang, M.; Yao, W.; Liu, J.; Niu, Q.; Chen, J.; Liu, Z.; Li, M.; Shi, B.; Pan, J. Functional role of lncRNA LOC101927497 in N-methyl-N'-nitro-N-nitrosoguanidine-induced malignantly transformed human gastric epithelial cells. *Life Sci.* **2018**, *193*, 93–103. [[CrossRef](#)] [[PubMed](#)]

45. Sánchez, Y.; Huarte, M. Long non-coding RNAs: Challenges for diagnosis and therapies. *Nucleic Acid Ther.* **2013**, *23*, 5–20. [[CrossRef](#)] [[PubMed](#)]
46. Liang, J.; Wei, X.; Liu, Z.; Cao, D.; Tang, Y.; Zou, Z.; Zhou, C.; Lu, Y. Long noncoding RNA CYTOR in cancer: A TCGA data review. *Clin. Chim. Acta* **2018**, *483*, 227–233. [[CrossRef](#)]
47. Zhang, X.; Tao, W. Long Noncoding RNA LINC00152 Facilitates the Leukemogenesis of Acute Myeloid Leukemia by Promoting CDK9 through miR-193a. *DNA Cell Biol.* **2019**, *38*, 236–242. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Underexpression of *LINC00173* in *TCF3/PBX1*-Positive Cases Is Associated With Poor Prognosis in Children With B-Cell Precursor Acute Lymphoblastic Leukemia

Didier Ismael May-Hau^{1,2}, Diego Alberto Bárcenas-López^{1,3}, Juan Carlos Núñez-Enríquez⁴, Vilma Carolina Bekker-Méndez⁵, Fredy Omar Beltrán-Anaya⁶, Elva Jiménez-Hernández⁷, Mónica Patricia Ortíz-Maganda⁵, Francisco Xavier Guerra-Castillo⁵, Aurora Medina-Sanson⁸, Janet Flores-Lujano⁴, Jorge Alfonso Martín-Trejo⁹, José Gabriel Peñaloza-González¹⁰, Martha Margarita Velázquez-Aviña¹⁰, José Refugio Torres-Nava¹¹, Gabriela Alicia Hernández-Echaurregui¹¹, Rosa Martha Espinosa-Elizondo¹², María de Lourdes Gutiérrez-Rivera¹³, Rodrigo Sanchez-Hernandez¹³, María Luisa Pérez-Saldívar⁴, Luz Victoria Flores-Villegas¹⁴, Laura Elizabeth Merino-Pasaye¹⁴, David Aldebarán Duarte-Rodríguez⁴, Minerva Mata-Rocha¹⁵, Omar Alejandro Sepúlveda-Robles¹⁵, Haydeé Rosas-Vargas¹⁵, Alfredo Hidalgo-Miranda¹, Juan Manuel Mejía-Arangur^{1,16*} and Silvia Jiménez-Morales^{1*}

OPEN ACCESS

Edited by:

Jignesh D. Dalal,
Case Western Reserve University,
United States

Reviewed by:

Divyaswathi Citla Sridhar,
University of Arkansas for Medical
Sciences, United States
Chintan Parekh,
University of Southern California,
United States
Jennifer Jarosca,
Medical University of South Carolina,
United States

*Correspondence:

Silvia Jiménez-Morales
sjimenez@inmegen.gob.mx
Juan Manuel Mejía-Arangur
jmejia@inmegen.gob.mx

Specialty section:

This article was submitted to
Hematologic Malignancies,
a section of the journal
Frontiers in Oncology

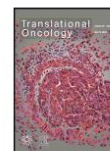
Received: 02 March 2022

Accepted: 21 April 2022

Published: 02 June 2022

¹ Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City, Mexico, ² Programa de Maestría en Investigación Clínica Experimental en Salud, Universidad Nacional Autónoma de México, Mexico City, Mexico, ³ Programa de Doctorado, Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, Mexico City, Mexico, ⁴ Unidad de Investigación Médica en Epidemiología Clínica, Hospital de Pediatría "Dr. Silvestre Frenk Freund", Centro Médico Nacional "Siglo XXI", Instituto Mexicano del Seguro Social, Mexico City, Mexico, ⁵ Unidad de Investigación Médica en Inmunología e Infectología, Hospital de Infectología "Dr. Daniel Méndez Hernández", Centro Médico Nacional "La Raza", Instituto Mexicano del Seguro Social, Mexico City, Mexico, ⁶ Laboratorio de Epidemiología Clínica y Molecular, Facultad de Ciencias Químico Biológicas, Universidad Autónoma de Guerrero, Chilpancingo, Mexico, ⁷ Servicio de Hematología Pediátrica, Hospital General "Gaudencio González Garza", Centro Médico Nacional "La Raza", Instituto Mexicano del Seguro Social, Mexico City, Mexico, ⁸ Departamento de Hemato-Oncología, Hospital Infantil de México Federico Gómez, Mexico City, Mexico, ⁹ Servicio de Hematología Pediátrica, Hospital de Pediatría "Dr. Silvestre Frenk Freund", Centro Médico Nacional "Siglo XXI", Instituto Mexicano del Seguro Social, Mexico City, Mexico, ¹⁰ Servicio de Onco-Pediatría, Hospital Juárez de México, Mexico City, Mexico, ¹¹ Servicio de Oncología, Hospital Pediátrico de Moctezuma, Secretaría de Salud de la Ciudad de México, Mexico City, Mexico, ¹² Servicio de Hematología Pediátrica, Hospital General de México, Mexico City, Mexico, ¹³ Servicio de Oncología Pediátrica, Hospital de Pediatría "Dr. Silvestre Frenk Freund", Centro Médico Nacional "Siglo XXI", Instituto Mexicano del Seguro Social, Mexico City, Mexico, ¹⁴ Servicio de Hematología Pediátrica, Centro Médico Nacional "20 de Noviembre", Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, Mexico City, Mexico, ¹⁵ Unidad de Investigación Médica en Genética Humana, Hospital de Pediatría "Dr. Silvestre Frenk Freund", Centro Médico Nacional "Siglo XXI", Instituto Mexicano del Seguro Social, Mexico City, Mexico, ¹⁶ Medicine Faculty, Universidad Autónoma de México, Mexico City, Mexico

Background: B-cell precursor acute lymphoblastic leukemia (BCP-ALL) is the most frequent pediatric cancer worldwide. Despite improvements in treatment regimens, approximately 20% of the cases cannot be cured, highlighting the necessity for identifying new biomarkers to improve the current clinical and molecular risk



Promising genes and variants to reduce chemotherapy adverse effects in acute lymphoblastic leukemia



Diego Alberto Bárcenas-López^{a,b}, Diana Karen Mendiola-Soto^{a,c}, Juan Carlos Núñez-Enríquez^d, Juan Manuel Mejía-Aranguré^{d,e}, Alfredo Hidalgo-Miranda^a, Silvia Jiménez-Morales^{a,*}

^a Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Periferico Sur 4809, Arenal Tepepan, Del. Tlalpan, Mexico City 14610, Mexico

^b Programa de Doctorado, Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, Mexico City, Mexico

^c Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México, Mexico City, Mexico

^d Unidad de Investigación Médica en Epidemiología Clínica, Hospital de Pediatría, CMNSXXI, Instituto Mexicano del Seguro Social, Mexico City, Mexico

^e Coordinación de Investigación en Salud, Instituto Mexicano del Seguro Social, Mexico City, Mexico

ARTICLE INFO

Keywords:

Acute lymphoblastic leukemia
Pharmacogenes
Polymorphisms
Treatment response
Personalized medicine

ABSTRACT

Almost two decades ago, the sequencing of the human genome and high throughput technologies came to revolutionize the clinical and therapeutic approaches of patients with complex human diseases. In acute lymphoblastic leukemia (ALL), the most frequent childhood malignancy, these technologies have enabled to characterize the genomic landscape of the disease and have significantly improved the survival rates of ALL patients. Despite this, adverse reactions from treatment such as toxicity, drug resistance and secondary tumors formation are still serious consequences of chemotherapy, and the main obstacles to reduce ALL-related mortality. It is well known that germline variants and somatic mutations in genes involved in drug metabolism impact the efficacy of drugs used in oncohematological diseases therapy. So far, a broader spectrum of clinically actionable alterations that seems to be crucial for the progression and treatment response have been identified. Although these results are promising, it is necessary to put this knowledge into the clinics to help physician make medical decisions and generate an impact in patients' health. This review summarizes the gene variants and clinically actionable mutations that modify the efficacy of antileukemic drugs. Therefore, knowing their genetic status before treatment is critical to reduce severe adverse effects, toxicities and life-threatening consequences in ALL patients.

Introduction

Acute lymphoblastic leukemia (ALL) is the commonest childhood cancer worldwide, accounting for approximately 80% of all childhood leukemias [1,2] and more than 349,000 newly cases were diagnosed in 2018 over the world [3]. ALL is a serious problem for developing countries, especially in Latin America, where the incidence of lymphoid leukemias is among the highest in the world and overall survival (OS) rates are lower compared to high-income countries (HIC) [4–7]. Currently, with risk-adapted stratification of patients, the development of targeted therapies in combination with conventional chemotherapy and the identification of novel molecular subtypes, OS in childhood ALL from HIC exceeds 90% at 5-years of follow-up but it is significantly lower in developing countries [5,8]. Furthermore, some patients relapse during early phases of therapy or develop resistance to treatment (either intrinsically or acquired resistance), and even those cases who are considered cured could suffer lifelong sequelae due to chemotherapy [9,10]. These survival rates disparities between racial groups are also

evident in developed countries, as it has been observed in Hispanics in the United States, where high-risk clinical features and lower outcomes are common [11,12]. It is well known that social and demographical aspects like low quality of care due to deficiencies in health infrastructure have a direct impact in mortality rates. Nonetheless, biological factors such as genetic variants and epigenetic processes are also determinant in the disease progression and in treatment response [7]. In fact, these aspects can also explain inter-individual and inter-population variations in the antileukemic drugs metabolism [11]. Identifying genes involved in drug response variability, also known as pharmacogenes, and detecting their genetic variation that influence both pharmacological efficacy and treatment toxicity, represents a big challenge [13].

The development of high throughput technologies in the last decade such as microarrays and next generation sequencing have allowed us to increase our knowledge about the genetic alterations that modulate drugs efficacy for many human diseases, particularly in cancer [14,15]. Due to their biological impact in the protein activity, studies of common germline variants as single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) in pharmacogenes, in addition to somatic

* Corresponding author.

E-mail address: sjimenez@inmegen.gob.mx (S. Jiménez-Morales).

<https://doi.org/10.1016/j.tranon.2020.100978>





Received 6 October 2020; Received in revised form 17 November 2020; Accepted 25 November 2020

1936-5233/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Review

Long Non-Coding RNA and Acute Leukemia

Gabriela Marisol Cruz-Miranda ¹, Alfredo Hidalgo-Miranda ², Diego Alberto Bárcenas-López ¹,
 Juan Carlos Núñez-Enríquez ³, Julian Ramirez-Bello ⁴, Juan Manuel Mejía-Aranguré ^{5,*}
 and Silvia Jiménez-Morales ^{2,*}

¹ Programa de Doctorado, Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; gmcm611@hotmail.com (G.M.C.-M.); d.a.barceñas@outlook.com (D.A.B.-L.)

² Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City 14610, Mexico; ahidalgo@inmegen.gob.mx

³ Unidad de Investigación Médica en Epidemiología Clínica, UMAE Hospital de Pediatría “Dr. Silvestre Frenk Freund”, Centro Médico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Mexico City 06720, Mexico; jcarlos_nu@hotmail.com

⁴ Unidad de Investigación en Enfermedades Metabólicas y Endócrinas, Hospital Juárez de México, Mexico City 07760, Mexico; dr.julian.ramirez.hjm@gmail.com

⁵ Coordinación de Investigación en Salud, Instituto Mexicano del Seguro Social, Mexico City 06720, Mexico

* Correspondence: juan.mejiaa@imss.gob.mx (J.M.M.-A.); sjimenez@inmegen.gob.mx (S.J.-M.); Tel.: +52-55-5350-1900 (ext. 1155) (S.J.-M.)

Received: 6 October 2018; Accepted: 22 October 2018; Published: 9 February 2019



Abstract: Acute leukemia (AL) is the main type of cancer in children worldwide. Mortality by this disease is high in developing countries and its etiology remains unanswered. Evidences showing the role of the long non-coding RNAs (lncRNAs) in the pathophysiology of hematological malignancies have increased drastically in the last decade. In addition to the contribution of these lncRNAs in leukemogenesis, recent studies have suggested that lncRNAs could be used as biomarkers in the diagnosis, prognosis, and therapeutic response in leukemia patients. The focus of this review is to describe the functional classification, biogenesis, and the role of lncRNAs in leukemogenesis, to summarize the evidence about the lncRNAs which are playing a role in AL, and how these genes could be useful as potential therapeutic targets.

Keywords: long non-coding RNAs; cancer; acute leukemia; therapeutic targets

1. Introduction

Leukemia is a group of hematological malignancies characterized by an oligoclonal expansion of abnormally differentiated, and sometimes poorly differentiated hematopoietic cells which infiltrate the bone marrow, and could also invade the blood and other extramedullary tissues. In general, AL can be divided into acute or chronic, and lymphoid or myeloid, according to their progression and affected lineage, respectively. Thus, we can identify the following subtypes: acute lymphoblastic leukemia (ALL), chronic lymphoblastic leukemia (CLL), acute myeloid leukemia (AML), and chronic myeloid leukemia (CML). AL is the main type of cancer in children worldwide [1,2]. In recent years, it has reported a trend of increase in the incidence AL; notwithstanding, the causes are still unclear. Studies conducted to identify the etiology of this disease have reported that a genetic background interacting with environmental factors (i.e., high doses of ionizing radiation, infections, parental occupational exposures, etc.) could explain this phenomenon [3]; however, the molecular mechanisms involved are not fully understood. To date, growing data have shown that different non-coding RNAs (ncRNAs) might be the link between the genome and the environment because they are closely related to normal physiological and pathological processes [4,5]. ncRNAs, also known

Tablas suplementarias

Tabla suplementaria 1. Términos empleados para el filtrado de fusiones

Término	Etiqueta	Descripción
mX (where X is a number)	--	count of pairs of reads supporting the fusion (excluding the mutimapping reads).
matched-normal	Banned	candidate fusion gene (which is supported by paired reads mapping on both genes and also by reads mapping on the junction point) was found also in the matched normal sample given as input to the command line option '--normal'
partial-matched-normal	Banned	candidate fusion gene (which is supported by paired reads mapping on both genes but no reads were found which map on the junction point) was found also in the matched normal sample given as input to the command line option '--normal'. This is much weaker than matched-normal.
short_repeats	Falso positivo	the sequence of the fusion junction contains a highly repetitive region containing repeating short sequences or polyA/C/G/T (detected using kmer = 2) . *A candidate fusion gene having this label has a good probability of being a false positive.
rt_circ_rna	Falso positivo	readthrough circular RNA
1000genomes	Falso positivo	fusion gene has been seen in a healthy sample. It has been found in RNA-seq data from some samples from 1000 genomes project. A candidate fusion gene having this label has a very high probability of being a false positive.
banned	Falso positivo	fusion gene is on a list of known false positive fusion genes. These were found with very strong supporting data in healthy samples (i.e. it showed up in file final-list_candidate_fusion_genes.txt). A candidate fusion gene having this label has a very high probability of being a false positive.
bodymap2	Falso positivo	fusion gene is on a list of known false positive fusion genes. It has been found in healthy human samples collected from 16 organs from Illumina BodyMap2 RNA-seq database. A candidate fusion gene having this label has a very high probability of being a false positive.
conjoining	Falso positivo	known conjoined genes (that is fusion genes found in samples from healthy patients) from the ConjoinG database (please use ConjoinG database for more information regarding the fusion gene). A candidate fusion gene having this label has a very high probability of being a false positive in case that one looks for fusion genes specific to a disease.
cortex	Falso positivo	fusion gene is on a list of known false positive fusion genes. It has been found in healthy human brains (BA9 prefrontal cortex) here. A candidate fusion gene having this label has a very high probability of being a false positive.
distance1000bp	Falso positivo	both genes are on the same strand and they are less than 1 000 bp apart. A candidate fusion gene having this label has a very high probability of being a false positive.
ensembl_fully_overlapping	Falso positivo	the genes forming the fusion gene are fully overlapping according to Ensembl database. A candidate fusion gene having this label has a very high probability of being a false positive.
ensembl_same_strand_overlapping	Falso positivo	the genes forming the fusion gene are fully/partially overlapping and are both on the same strand according to Ensembl database. *A candidate fusion gene having this label has a very high probability of being a false positive (this is most likely and alternative splicing event).
gtex	Falso positivo	fusion gene has been seen in a healthy sample. It has been found in GTEx database of healthy tissues (thru FusionAnnotator). A candidate fusion gene having this label has a very high probability of being a false positive.
hpa	Falso positivo	fusion gene has been seen in a healthy sample. It has been found in RNA-seq database of 27 healthy tissues. A candidate fusion gene having this label has a very high probability of being a false positive.
mt	Falso positivo	one or both genes are situated on mitochondrion. A candidate fusion gene having this label has a very high probability of being a false positive.
non_cancer_tissues	Falso positivo	fusion gene which has been previously reported/found in non-cancer tissues and cell lines in Babiceanu et al, Recurrent chimeric fusion RNAs in non-cancer tissues and cells, Nucl. Acids Res. 2016. These are considered as non-somatic mutation and therefore they may be skipped and not reported.
non_tumor_cells	Falso positivo	fusion gene which has been previously reported/found in non-tumor cell lines, like for example HEK293. These are considered as non-somatic mutation and therefore may be skipped and not reported.

pair_pseudo_genes	Falso positivo	one gene is the other's pseudogene. A candidate fusion gene having this label has a very high probability of being a false positive.
paralogs	Falso positivo	both genes involved in the fusion gene are paralog for each other (most likely this is a false positive fusion gene). A candidate fusion gene having this label has a very high probability of being a false positive.
refseq_fully_overlapping	Falso positivo	the genes forming the fusion gene are fully overlapping according to RefSeq NCBI database. A candidate fusion gene having this label has a very high probability of being a false positive.
refseq_same_strand_overlapping	Falso positivo	the genes forming the fusion gene are fully/partially overlapping and are both on the same strand according to RefSeq NCBI database. *A candidate fusion gene having this label has a very high probability of being a false positive (this is most likely and alternative splicing event).
ribosomal	Falso positivo	one or both gene is a gene encoding for ribosomal protein
rna	Falso positivo	one or both genes is a rRNA. A candidate fusion gene having this label has a very high probability of being a false positive.
similar_symbols	Falso positivo	both genes have the same or very similar gene names (for example: RP11ADF.1 and RP11ADF.2). A candidate fusion gene having this label has a very high probability of being a false positive.
snorna	Falso positivo	one or both genes is a snoRNA
snrna	Falso positivo	one or both genes is a snRNA
trna	Falso positivo	one or both genes is a tRNA
ucsc_fully_overlapping	Falso positivo	the genes forming the fusion gene are fully overlapping according to UCSC database. A candidate fusion gene having this label has a very high probability of being a false positive.
ucsc_same_strand_overlapping	Falso positivo	the genes forming the fusion gene are fully/partially overlapping and are both on the same strand according to UCSC database. *A candidate fusion gene having this label has a very high probability of being a false positive (this is most likely and alternative splicing event).
cacg	Falso positivo	known conjoined genes (that is fusion genes found in samples from healthy patients) from the CACG database (please see CACG database for more information). A candidate fusion gene having this label has a very high probability of being a false positive in case that one looks for fusion genes specific to a disease.
pseudogene	Falso positivo	one or both of the genes is a pseudogene
tcga-normal	Falso positivo	known fusion gene from healthy samples from TCGA database (please use Google for more information regarding the fusion gene)
readthrough	Pendiente	the fusion gene is a readthrough event (that is both genes forming the fusion are on the same strand and there is no known gene situated in between); Please notice, that many of readthrough fusion genes might be false positive fusion genes due to errors in Ensembl database annotation (for example, one gene is annotated in Ensembl database as two separate genes). A candidate fusion gene having this label has a high probability of being a false positive.
adjacent	Pendiente	both genes forming the fusion are adjacent on the genome (i.e. same strand and there is no other genes situated between them on the same strand)
distance100kbp	Pendiente	both genes are on the same strand and they are less than 100 000 bp apart. A candidate fusion gene having this label has a higher probability than expected of being a false positive.
distance10kbp	Pendiente	both genes are on the same strand and they are less than 10 000 bp apart. A candidate fusion gene having this label has a higher probability than expected of being a false positive.
duplicates	Pendiente	both genes involved in the fusion gene are paralog for each other. For more see Duplicated Genes Database (DGD) database . A candidate fusion gene having this label has a higher probability than expected of being a false positive.
ensembl_partially_overlapping	Pendiente	the genes forming the fusion gene are partially overlapping (on same strand or on different strands) according the Ensembl database. *A candidate fusion gene having this label has a good probability of being a false positive.
fragments	Pendiente	the genes forming the fusion are supported by only and only one fragment of RNA. A candidate fusion gene having this label has a medium probability of being a false positive.
healthy	Pendiente	fusion gene has been seen in a healthy sample. These have been found in healthy samples but the support for them is less strong (i.e. paired reads were found to map on

		both genes but no fusion junction was found) than in the case of banned label (i.e. it showed up in file preliminary list of candidate fusion genes). Also genes which have some degree of sequence similarity may show up marked like this. A candidate fusion gene having this label has a small probability of being a false positive in case that one looks for fusion genes specific to a disease.
lincrna	Pendiente	one or both genes is a lincRNA
metazoa	Pendiente	one or both genes is a metazoa_srp gene Metazia_srp
mirna	Pendiente	one or both genes is a miRNA
similar_reads	Pendiente	both genes have the same reads which map simultaneously on both of them (this is an indicator of how similar are the sequences of both genes; ideally this should be zero or as close to zero as possible for a real fusion). A candidate fusion gene having this label has a very high probability of being a false positive.
ucsc_partially_overlapping	Pendiente	the genes forming the fusion gene are partially overlapping (on same strand or on different strands) according the UCSC database. *A candidate fusion gene having this label has a good probability of being a false positive.
yrna	Pendiente	one or both genes is a Y RNA
antisense	Pendiente	one or both genes is a gene coding for antisense RNA
m0	Pendiente	There are no pairs of non-multi-mapping reads supporting the fusion. Basically, there are supporting pairs of reads but all of them map also in some other places on genome (that is their mappings on genome are not unique).
no_protein	Pendiente	one or both genes have no known protein product
long_repeats	Pendiente	the sequence of the fusion junction contains a highly repetitive region containing repeating long sequences or polyA/C/G/T (detected using kmer = 9) . *A candidate fusion gene having this label has a good probability of being a false positive.
multi	Pendiente	one of the genes of both have multi-mapping reads mapping (which map simultaneously also on other gene/genes
refseq_partially_overlapping	Pendiente	the genes forming the fusion gene are partially overlapping (on same strand or on different strands) according the RefSeq NCBI. *A candidate fusion gene having this label has a good probability of being a false positive.
short_distance	Pendiente	both genes are on the same strand and they are less than X bp apart, where X is set using the option '--dist-fusion' and by default it is 200 000 bp. A candidate fusion gene having this label has a higher probability than expected of being a false positive.
18cancers	Probable fusión	fusion gene found in a RNA-seq dataset of 18 types of cancers from 600 tumor samples published here.
cell_lines	Probable fusión	known fusion gene from paper: C. Klijn et al., A comprehensive transcriptional portrait of human cancer cell lines, Nature Biotechnology, Dec. 2014, DOI:10.1038/nbt.3080
ccl	Probable fusión	known fusion gene found in human cancer cell lines that are in CCLE (Cancer Cell Line Encyclopedia); known fusions from DepMap portal
ccl3	Probable fusión	known fusion gene found in human cancer cell lines that are in CCLE (Cancer Cell Line Encyclopedia); known fusions from from paper: Vellichirammal et al., Pan-Cancer Analysis Reveals the Diverse Landscape of Novel Sense and Antisense Fusion Transcripts, Mol. Ther. Nucleic Acids, 2020, DOI:10.1016/j.omtn.2020.01.023
cgp	Probable fusión	known fusion gene from the CGP database
chimerdb2	Probable fusión	known fusion gene from the ChimerDB 2 database
chimerdb3kb	Probable fusión	known fusion gene from the ChimerDB 3 KB (literature curation) database
chimerdb3pub	Probable fusión	known fusion gene from the ChimerDB 3 PUB (PubMed articles) database
chimerdb3seq	Probable fusión	known fusion gene from the ChimerDB 3 SEQ (TCGA) database
cosmic	Probable fusión	known fusion gene from the COSMIC database (please use COSMIC database for more information regarding the fusion gene)
known	Probable fusión	fusion gene which has been previously reported or published in scientific articles/reports/books/abstracts/databases indexed by Google, Google Scholar, PubMed, etc. This label has only the role to answer with YES or NO the question "has ever before a given (candidate) fusion gene been published or reported?". This label does not have in anyway the role to provide the original references to the original scientific articles/reports/books/abstracts/databases for a given fusion gene.
oesophagus	Probable fusión	fusion gene found in a oesophageal tumors from TCGA samples, which are published here.

oncogene	Probable fusión	one gene or both genes are a known oncogene according to ONGENE database
cancer	Probable fusión	one gene or both genes are cancer associated according to Cancer Gene database
tumor	Probable fusión	one gene or both genes are proto-oncogene or tumor suppressor gene according to UniProt database
pancreases	Probable fusión	known fusion gene found in pancreatic tumors from article: P. Bailey et al., Genomic analyses identify molecular subtypes of pancreatic cancer, Nature, Feb. 2016, http://dx.doi.org/110.1038/nature16965
prostates	Probable fusión	known fusion gene found in 150 prostate tumors RNAs from paper: D. Robison et al, Integrative Clinical Genomics of Advanced Prostate Cancer, Cell, Vol. 161, May 2015, http://dx.doi.org/10.1016/j.cell.2015.05.001
tcga	Probable fusión	known fusion gene from the TCGA database (please use Google for more information regarding the fusion gene)
tcga-cancer	Probable fusión	known fusion gene from tumor samples from TCGA database (please use Google for more information regarding the fusion gene)
tcga2	Probable fusión	known fusion gene from the TCGA database (please use Google for more information regarding the fusion gene)
tcga3	Probable fusión	known fusion gene from the TCGA database (please use Google for more information regarding the fusion gene)
exon-exon	Probable fusión	the fusion junction point is exactly at the known exon's borders of both genes forming the candidate fusion
gliomas	Probable fusión	fusion gene found in a RNA-seq dataset of 272 glioblastoms published here.
ticdb	Probable fusión	known fusion gene from the TICdb database (please use TICdb database for more information regarding the fusion gene)
Fitrado manual de términos para identificar falsos positivos		
1. La misma muestra muestra varios eventos de fusion (diferentes partners) con un mismo gen (gen "promiscuo")		
2. Fusiones que involucran a genes de la hemoglobina (ej. HBA, HBB)		
3. Fusiones que involucran genes estructurales o de expresión constitutiva (ej. ACTB, GAPDH, B2M)		
4. Fusiones que involucran genes ribosomales (rRNA)		
5. Fusiones que involucran genes del HLA		
6. Genes altamente expresados (abajo descripción). Algunas fusiones son específicas de un lote de secuenciación, de un algoritmo, o de un tipo celular. El gen WDR74 se encuentra frecuentemente referido como candidato por estar asociado a cáncer en la literatura, pero está altamente expresado en precursores linfoides		
7. Fusiones que no tengan una etiqueta en descripción (no se puede inferir su papel biológico, o no existe información en bases de datos) o que tengan pocos términos en la columna "Fusion description".		
**** Genes altamente expresados tienen mayor probabilidad de aparecer como falso positivo (ej. WDR74, es altamente expresado en MO y tejidos linfoides, ver https://www.proteinatlas.org/ENSG00000133316-WDR74/tissue)		
* Del programa Arriba: <i>True predictions usually have a balanced number of split reads and discordant mates. Events with only discordant mates or without discordant mates and only split reads having anchors in just one gene are frequently artifacts.</i>		
** Del programa Arriba:	<i>Most of the time, these events have very few split-reads (1-2) while having a high number of discordant mates (dozens).</i>	
	<i>Often, multiple genes form a network of in vitro fusions. The output file then contains many events between a small set of genes. These genes are typically expressed at a high level (in the top percentile).</i>	

Tabla suplementaria 2. Secuencia y conteos de lecturas de fusiones identificadas en LLA

ID	Gen 5'	Gen 3'	# PR	# SR	Método	Secuencia de fusión	Efecto predicho	¿Alteración conductora o concomitante?	Subtipo asignado	Comentario adicional
61	BCR	ABL1	111	27	FC	CGGGCAGATCTGGCCCAACGATGGCGAGGGCGCCTTCATGGAGACGCAG*AAGCCCTCAGCGGCCAGTAGCATCTGACTTTGAGCCTCAGGGTCTGAGT	En marco de lectura	Conductora	BCR-ABL1	Fusión recíproca
61	IGH@	CCND3	3	2	FC	TATTGCACTACTGCACTCCAGCCTGGGTGACAGAGTGAGACTCCATCTCA*AAACAAACAAACAAACAAACAAACAAATACAGAAACATTATAAAATCACAGA	---/intrónica	Concomitante	BCR-ABL1	
84	IGH@	CRLF2	5	2	FC	CCTGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCGGGTACGTGGGATTAC*AGCAGAAGGAGTACAGATTCAGATCATCTACTTCAATTTAGAAACCGTGC	---/intrónica	Concomitante	CRLF2	
84	P2RY8	CRLF2	144	48	FC	GGAGACCCTCTGAGCTCTCACCTGCTACTCTGCGCGTCTCTGACAG*GCATGGGGCGGTGTTCTGCTGTGGGGAGCTGCGCTTTCTGCTGGGA	UTR/UTR	Conductora	CRLF2	
85	TCF3	ZNF384	8	14	FC	TCCAGCCACACGCGCCTGTCAGCGGGGGCCGACAGCCTCC TGG*AAAGTTCAGGAGCCCTGGAAAGGAGAAGGAATAAGACGGCAGG	Secuencia codificante trunca/UTR	Conductora	ZNF384	
143	DUX4	IGH@	62	48	FC	GGCGCCCTCCAGCGCTCAGGAGCCGCGCCCTGGTCTGCATCCC*CTGAGAAAATGGTTAAGAAAATTATGACTTAAAATGTGAGAGATTTTC	Secuencia codificante trunca/---	Conductora	DUX4	
190	CRLF2	IGH	16	13	Búsqueda dirigida	AGTGGGTAGATAAATGATAGACATAGATAGACAGATATAG GTTCCACGGGTAGTGGGGTACTACTACTACGGTATGGACGTCTGGGGCCAAGGGACCA	Intergénica/---	Conductora	CRLF2	
203	TPM4	KLF2	6	12	FC	CCTGCAGCGGGAGCTGGACGGCGAGCGGAGCGGCGCGGAGAAA*GTGAGAAGCCCTACCATGCAACTGGGACGGCTGCGGCTGGAA	En marco de lectura	No se conoce	Otro	
205	TCF3	PBX1	249	37	FC	CCAGCCAGGCACCCTCCCTGACCTGTCTGGCCTCCGACTCCTACAGTG*TTTTGAGTATCCGAGGAGCCAGGAGGAGGAACCCACAGACCCCAAGCTG	En marco de lectura	Conductora	TCF3-PBX1	
207	CCND2	TRA@	6	3	FC	TCTCTCTCTCTCTCAGTTATGTAGTTTCTTGTCTGGACTTTTTTT*TCTTTTCTTTTTCTTTTTTTTCTTCTCTGTTTCAGATAGTCTCACTCT	UTR/---	No se conoce	Concomitante	Fusión recíproca
207	MELK	RNF38	3	2	FC	AGCACCTGAGTTAATACAAGGCAAATCATATCTTGGATCAGAG*ATATCTCCGGGGCAATTCAGCATCTACTACGGCCATCCTA	En marco de lectura	No se conoce	Concomitante	
271	PAX5	NOL4L	70	21	FC	CTCAGACATCTCACCACCACAGAGCCATCAAGCCCGAGCAG*TCACCCCTACAGCTCTGGGAGTACGATTCCATCAAGACCG	En marco de lectura	Conductora	PAX5	
414	MEF2D	SS18	45	23	FC	GGCTGAGGGCTCAGCTGTGAAGAGGATGCGGCTTGATACCTGG*CCACCACACAGAATATGCCTATGGTCTGGAGGATGAATC	En marco de lectura	Conductora	MEF2D	Dos isoformas
475	NCL	BAZ2A	5	2	FC	TCCTCCTCAAAGGAGGTAGAAGAAGATAGTGAAGATGAGGAAATGTCAG*AAGAAAAGGAAAGACAAAGCAGGAAAAACTGAAGAAAAAGTCAAGAGG	En marco de lectura	No se conoce	Otro	
475	PLPBP	SET	3	4	FC	TTCATTGGCCACCTACAGAACAATAATGTCAACAATTGATGG*AAAAAGAACAGCAAGAAGCGATTGAACACATTGATGAAGTACA	En marco de lectura	No se conoce	Otro	
528	CCND3	IGH@	3	2	FC	CACTCCAGCCTGGGGACAGAGTGAGACTTCTTCTCAAAAATTAATAAAT*ATGGTGGCAGGCACCTGTAATCCAGCTACTGGGAGGCTGACTGGAG	intrónica/---	No se conoce	Otro	
528	SMARCA4	CARM1	5	4	FC	CCGGTCCCGCCCGCCACCTCCAAATTACAGCAGGCCTCATG*ATGAAGATGTGTGTCTTTAAGTGTCTAGTGTCCCGAGAGAC	En marco de lectura	No se conoce	Otro	
540	KDM4C	PLEKHO1	4	3	FC	CCTCCTCCAGGAGTCTGCTCTCGCCCAACCCGCGGCCAG*GGACCTCAGGATGGAAACCAGCAGCCTGCACCCGCCAGAAAG	UTR/Secuencia codificante trunca	No se conoce	Otro	
540	RAF1	TMEM40	6	6	FC	TCCAGCTAATTGACATTGCCCGGCAGAGGCTCAGGGAATGGA*GAAAAGCCATGGAGACTTCAGATCCTCTCCAGCCTCAGGA	Secuencia codificante trunca/UTR	No se conoce	Otro	Fusión recíproca
546	TCF3	PBX1	92	24	FC	GGCACCCTCCGTACCTGTCTCGCCTCCGACTCCTACAGTG*TTTTGAGTATCCGAGGAGCCAGGAGGAAACCCACAGACCC	En marco de lectura	Conductora	TCF3-PBX1	
600	IGH@	BCL2	3	3	FC	GAACTTGGAGATCCCAAGTGGGCACCAATGACTTACTGTTTACTGACTT*CCCCAACTCAGATTTAATAAATATTGACATTA GCCACATTTTAAAAATAT	---/intrónica	Concomitante	PAX5	
600	PAX5	ST18	70	21	FC	ACGTCCCCAGCGCCGACACCAACAGCGCAAGAGAGAGCAAG*AGGAATCATGATAAACCTATGGATGAAAGTCTTCTTTCAAC	En marco de lectura	Conductora	PAX5	

Tabla suplementaria 3. Secuencia de oligonucleótidos para validación de fusiones génicas

Fusión	Primer Forward (5' - 3')	Primer Reverse (5' - 3')	Tamaño amplicón	Validación
ACTB	CCTGTGGCATCCACGAAACT	GGGGCCGGACTCGTCATAC	290	Gen control
B2M	GAGTATGCCTGCCGTGTGAA	GCTTACATGTCTCGATCCCACT	74	Gen control
BCR-ABL1 mayor	TCACTGCTGCTGCTTATGTCT	GGCGTGATGATGTTGCTTGG	515	RT-PCR
BCR-ABL1 menor	ATTGCGGAGGCGGCTATAC	ACCCGGAGCTTTTCACCTTT	502	RT-PCR
CDKN2A-TRDC	CGAGTGAGGGTTTTCTGGT	GCAGGTACCGTGCACATC	380	RT-PCR y Sanger
CRLF2-IGH	GAAGAATGAGAGGCGGGTAA	GACGGGAATTCTCACAGGA	206	RT-PCR y Sanger
DUX4-ETV6	GGGCAACCTCTCCTAGAAAC	GCGGATCGAGTCTTCTCCA	203	RT-PCR y Sanger
ETV6-RUNX1	CACCAGGAGTCTACCCTCT	ATGGTGGGTAGCATTCT	719	RT-PCR
MEF2D-HNRNPUL1	TGGGGTCTCCAGTCTACTC	TCACTAGCTTGCCGCTTCC	200	RT-PCR
MEF2D-SS18	TGGGGTCTCCAGTCTACTC	TCTGGCCGTTTCATCTGGTTC	809	RT-PCR
NCL-BAZ2A	GTTCAACGGGAGAGTGTCA	TCTGTGTGGCCAGGGTCTTA	246	Sin RNA/cDNA disponible
P2RY8-CRLF2	ACTTTGAGTGACCACGGCTT	AGTCAGGTTGGTCTGGAGT	320	RT-PCR
PAX5-NOL4L	GGACAAAAGTACAGCAGCCAC	TGGTCATCGTGGTTCATCGTC	484	RT-PCR
PAX5-ST18	GGACAAAAGTACAGCAGCCAC	CACTCGTCTGCTTCATCGCTC	409	RT-PCR y Sanger
PLBP-SET	GTGATCGAGGCCTATGGACA	TGGTTGGCGGAGTTTGTAT	285	Sin RNA/cDNA disponible
TCF3-PBX1	ACCAGCCTCATGCACAACCA	CAGAAATCGGGAACGCAGGA	513	RT-PCR
TCF3-ZNF384	ATGGAGCAGAGGTGAACGGT	ACTGTCAGCAAGGTGGGGTA	349	RT-PCR y Sanger

Tabla suplementaria 4. Fármacos quimioterapéuticos empleados en el tratamiento de la LLA

	Inducción a la remisión (4-6 semanas)	Consolidación (~8 semanas)	Mantenimiento (~2 a 3 años)
Centro Médico Siglo XXI HP/7/08 (Basado en el protocolo ALL-BFM 95)	-Prednisona -Vincristina -Metotrexato (IT) -Daunorrubicina -L-asparaginasa RA: -Radioterapia (12 Gy)	-Ciclofosfamida -Citarabina -Metotrexato (IT) -Mercaptopurina RA: -Citarabina -L-asparaginasa -Vincristina -Metotrexato -Daunorrubicina	-Mercaptopurina (1 años más en varones) -Metotrexato RA: -Vincristina y esteroides (en pulsos)
Centro Médico "La Raza" DFCI 00-01	-Prednisona -Vincristina -Metotrexato (IT) -Daunorrubicina + dexrazoxano -L-asparaginasa RA: -Radioterapia (12-18 Gy)	-Vincristina -L-asparaginasa -Prednisona -Mercaptopurina -Metotrexato (IT) RH: -Metotrexato (IM)	-Vincristina + metotrexato + Prednisona (en pulsos) -Mercaptopurina

		RA: - Doxorubicina (IV) + dexrazoxano	
-Hospital Pediátrico Moctezuma -Hospital Juárez de México -Hospital Infantil de México -Centro Médico Nacional "20 de Noviembre" St Jude XV o Protocolo del Seguro Popular (Basado en el St Jude XV)	-Prednisona -Vincristina -L-asparaginasa -Metotrexate (IT) -Ciclofosfamida -Citarabina -Daunorrubicina	-Metotrexato -Mercaptopurina -Citarabina -L-asparaginasa -Metotrexato (IT)	-Mercaptopurina -Metotrexato RH: -Dexametasona -Vincristina (pulsos) RA: -L-asparaginasa -Dexametasona -Vincristina -Doxorrubicina

IT= Intratecal; IM= Intramuscular; IV= Intravenosa; Gy= Grays; RH= Riesgo habitual; RA= Riesgo alto.

Las dosis, la duración de cada fase, y combinación de fármacos varían en función del riesgo del paciente y el esquema.

Tomados y modificados de: Möricke et al., 2008; Pui et al., 2009; Jiménez Hernández et al., 2015; Omar et al., 2022.

Tabla suplementaria 5. Características de los pacientes que presentaron recaída temprana (RT)

n = 23

	Total n= 23 n (%)	LLA-RT y muerte n= 16 n (%)	LLA-RT vivos n= 7 n (%)	P*
Sexo				
Masculino	13 (56.5)	8 (50)	5 (71.4)	.4
Femenino	10 (43.5)	8 (50)	2 (28.6)	
Edad (años)				
1 – 9.99	9 (39.1)	4 (25)	5 (71.4)	.06
≥ 10	14 (60.9)	12 (75)	2 (28.6)	
CL. (x10⁹ / L)				
<10	7 (30.4)	5 (31.25)	2 (28.8)	.96
10-49.99	7 (30.4)	5 (31.25)	2 (28.6)	
50-99.99	4 (17.4)	3 (18.75)	1 (14.2)	
≥ 100	5 (21.7)	3 (18.75)	2 (28.6)	
Inmunofenotipo				
Linaje B	21 (91.3)	15 (93.75)	6 (85.8)	.52
Linaje T	2 (8.7)	1 (6.25)	1 (14.2)	
Esquema de quimio.				
St Jude Total XV	16 (69.6)	10 (62.5)	6 (85.8)	-
St Jude Total XIII y XV	2 (8.7)	2 (12.5)	0 (0)	

DFCI-00-01	1 (4.3)	1 (6.25)	0 (0)	
BFM-95	4 (17.4)	3 (18.75)	1 (14.2)	
Sitio de recaída				
MO aislada	15 (65.2)	13 (81.25)	2 (28.6)	.04
SNC	5 (21.7)	2 (12.5)	3 (42.8)	
Combinada***	3 (13.1)	1 (6.25)	2 (28.6)	

*Valor de p de comparación RT y muerte vs RT sin muerte (prueba de χ^2 o exacta de Fisher según el caso)

**Combinada= Recaída en MO aislada + SNC y/o testículos

CL = Cuenta leucocitaria en sangre periférica

MO= Médula ósea

SNC= Sistema nervioso central

Congresos internacionales

El presente trabajo fue aceptado como póster en:

- International Conference on 2nd Translational Research E- Conference: Acute Lymphoblastic Leukaemia. European School of Haematology, Mayo de 2021
- 64th ASH Annual Meeting and Exposition. American Society of Hematology. Nueva Orleans, Louisiana, Diciembre de 2022.



May 7th, 2021

CERTIFICATE OF ATTENDANCE

This is to certify that

Diego Alberto Barcenás-Lopez

attended the International Conference on 2nd Translational Research E- Conference: Acute Lymphoblastic Leukaemia

which took place from May 7 to 9, 2021

under the scientific direction of

Jan Cools, Hervé Dombret, Adele Fielding, Mark Litzow

EUROPEAN SCHOOL OF HAEMATOLOGY (ESH)
 IJH, Centre Hayem
 Hôpital Saint-Louis
 1, Av. Claude Vellefaux
 75010 PARIS, FRANCE

Whole Transcriptome Analysis in Mexican Children with High-Risk Acute Lymphoblastic Leukemia: A Preliminary Report from the Migiccl

Diego Alberto Barcenás-López, Juan Carlos Nuñez-Enríquez, Laura Lucila Gomez-Romero, Hugo Antonio Tovar-Romero, Elva Jimenez-Hernandez, Aurora Medina-Sanson, Jose Gabriel Peñalosa-Gonzalez, Luz Victoria Flores Villegas, Minerva Mata-Rocha, Haydee Rosas-Vargas, Alfredo Hidalgo-Miranda, Juan Manuel Mejia-Arangure, Silvia Jimenez-Morales



Blood (2022) 140 (Supplement 1): 6365-6366.

<https://doi.org/10.1182/blood-2022-168167>



Abstract

Almost 20% of Mexican children with acute lymphoblastic leukemia (ALL) die within the first 2 years, and relapses are over 4 times more frequent in Mexico than in high-income countries, mainly early relapses (<36 months from diagnosis or at treatment completion)(Nuñez-Enríquez et al., BMC Cancer 2019). It has been observed that gene fusions associated with good prognosis are less common in Mexican ALL cases, but biomarkers associated with older age and poor prognostic molecular subtypes are more frequent. These observations have been reported in Hispanics in the US and in other Latin-American populations. For instance, *ETV6-RUNX1*, a favorable prognosis marker, shows low frequency in Mexican children and other Native American populations (~10%), but Ph-like subtype is more common in Hispanics (11-15%) and has been associated with ancestry (Lee et al., JAMA Oncol, 2022). This project aimed to identify gene fusions by whole transcriptome sequencing (RNA-seq) in a sub-cohort of Mexican children with ALL.

We carried out a retrospective multicenter cohort study as part of the Mexican Interinstitutional Group for the Identification of the Causes of Childhood Leukemia (MIGICCL). From a cohort of 496 children with ALL treated in public hospitals from Mexico City, we selected 49 patients for RNA-seq analysis. All patients were followed-up for at least 36 months. Total RNA was extracted from bone marrow and sequencing was done with Illumina. Fusion detection was performed with FusionCatcher software and differential expression was done with DESeq2. All fusions were manually curated and inspected in IGV software to eliminate false positives. In cases without an identified fusion, raw reads were aligned to GRCh38 genome with STAR software and soft-clipped reads mapping upstream of *CRLF2*, *DUX4*, and to *IGH* locus were searched to identify *IGH* alterations. Known fusions were validated by RT-PCR and novel ones by RT-PCR and Sanger sequencing.

By National Cancer Institute (NCI, USA) risk criteria, thirty-five (71.4%) patients were classified as high-risk, 23 (47%) had early relapse (Table 1). The 4-year overall survival (OS) of the studied population was 50% (95% CI: 37-67). No differences in EFS and OS were observed between patients classified as high- or standard-risk ($p=0.87$ and $p=0.78$, respectively). Overall, 65.3% of patients had at least one fusion. Thirty-one different fusions were found and 14 are considered recurrent in B-ALL. Nine fusions (18.4%) were not previously reported or had a novel partner fused to a recurrently mutated gene. By grouping them into well-known B-ALL subtypes, the most common were *DUX4* (13%), *CRLF2* (10.87%) and *BCR-ABL1* (10.87%) (Figure 1). Four *DUX4* cases had *DUX4-IGH* fusion, one case had a novel *DUX4-ETV6* fusion and in another one the underlying genetic alteration could not be found but had the highest *DUX4* expression. Regarding *CRLF2*, *P2RY8-CRLF2* was the most common fusion (4/5 cases), and one case had a *CRLF2-IGH* fusion that was missed by the fusion detection software but identified with targeted search via soft-clipped reads. *ETV6-RUNX1* was found in 6.5%, in line with the low frequency reported in Latinos. The frequency of newly discovered subtypes like *MEF2D*, *PAX5* (including a novel 3' partner in *PAX5-ST18*) and *ZNF384* alterations was 6.5%, 4.3% and 2.2%, respectively, similar frequencies to those reported in bigger cohorts. As expected, none of *TCF3-PBX1* and *ETV6-RUNX1* patients died during follow-up. The remaining genetic subtypes had poor outcome. Four-year OS for kinase-driven subtypes was 50%, *P2RY8-CRLF2* having the worst outcome (3/4 died before treatment completion). It was surprising to find the low OS of *DUX4*, a subtype associated with an excellent prognosis even in adult ALL. The 4-year OS of *DUX4* was 33%, and all patients classified as *DUX4* who relapsed also died during treatment. The outcome of patients without fusions was also poor, of just 51% after 4 years.

We report the first RNA-seq-based analysis in Mexican children with high-risk ALL. We were able to classify around 80% of patients based on transcriptome data alone. Although the implementation of these technologies in a clinical setting is still a challenge, especially for developing countries, some genetic subtypes deserve consideration to explore their prognostic significance in bigger cohorts, perhaps using more feasible techniques to identify them, and to provide a more precise risk stratification.

Figure 1

