



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**DOCTORADO EN CIENCIAS BIOMÉDICAS**

**PERFIL DE EXPRESIÓN GLOBAL DE LINFOBLASTOS PRE B  
CD10+/CD19+ DE PACIENTES CON LEUCEMIA AGUDA  
LINFOBLÁSTICA B (LAL-B) Y ANÁLISIS COMPARATIVO CON  
PACIENTES HISPANOS CON LAL-B DE LA BASE DE DATOS  
TARGET**

**TESIS**

**QUE PARA OPTAR POR EL GRADO DE:  
DOCTORA EN CIENCIAS**

**PRESENTA:**

**MA. DE JESÚS LAURA CASTAÑEDA PARTIDA**

**TUTOR PRINCIPAL:**

**DR. JUAN BAUTISTA PATRICIO GARIGLIO VIDAL  
INSTITUTO DE FISIOLÓGÍA CELULAR**

Cd. Universitaria, Ciudad de México, Septiembre 2022.



Universidad Nacional  
Autónoma de México



**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.

# ÍNDICE

<b>RESUMEN.....</b>	<b>5</b>
<b>ABSTRACT.....</b>	<b>7</b>
<b>INTRODUCCIÓN.....</b>	<b>9</b>
<b>Justificación.....</b>	<b>10</b>
<b>Objetivos.....</b>	<b>11</b>
<b>Hipotesis.....</b>	<b>12</b>
<b>MATERIAL Y MÉTODOS.....</b>	<b>12</b>
<b>Pacientes.....</b>	<b>12</b>
<b>Extracción de células mononucleares.....</b>	<b>13</b>
<b>Immunotipificación y extracción de linfoblastos pre-B CD10+/CD19+ mediante</b>	
<b>FACS (Fluorescence Activated Cell</b>	
<b>Sorting).....</b>	<b>13</b>
<b>Aislamiento de RNA total.....</b>	<b>14</b>
<b>Análisis de Microarreglos.....</b>	<b>15</b>
<b>Análisis bioinformático.....</b>	<b>15</b>
<b>Análisis de tiempo real.....</b>	<b>16</b>
<b>Análisis bioinformático comparativo con muestras de pacientes Hispanos</b>	
<b>disgnosticados con LAL-B de la base de datos pública <i>TARGET</i></b>	
<b>(Therapeutically Applicable Research to Generate Effective Treatments).....</b>	<b>17</b>
<b>RESULTADOS.....</b>	<b>18</b>
<b>DISCUSIÓN.....</b>	<b>37</b>
<b>CONCLUSIONES.....</b>	<b>45</b>

<b>APÉNDICE DE FIGURAS SUPLEMENTARIAS.....</b>	<b>47</b>
<b>APENDICE DE PROTOCOLOS DE TRABAJO.....</b>	<b>51</b>
<b>REFERENCIAS.....</b>	<b>62</b>

***“Porque la vida de la carne en la sangre está...”***

***Levítico 17:11***

## RESUMEN

La Leucemia Aguda Linfoblástica (LAL) es el cáncer infantil más frecuente. La Ciudad de México tiene una de las incidencias más altas de LAL-B en el mundo con pacientes que muestran una baja respuesta a la terapia convencional, baja tasa de sobrevivencia y altas tasas de recaída. Para estudiar más las características moleculares de la LAL-B en niños mexicanos, se implementó un protocolo por FACS (Flourescence Activated Cell Sorting) para aislar linfoblastos precursores B (pre-B) CD10+/CD19+ de nueve muestras de sangre periférica (SP) y cuatro aspirados de médula ósea (MO) de pacientes diagnosticados con LAL-B. Se llevó a cabo un perfil de expresión génica global y un análisis jerárquico de cluster que reveló 136 genes diferencialmente expresados (GDEs): 62 sobreexpresados (45.6%) y 74 subexpresados (54.4%). Se calculó el coeficiente de correlación de Pearson para encontrar similitud entre las poblaciones de linfoblastos pre-B de SP y MO. Se seleccionaron 26 de los principales GDEs y se validaron 21 por RT-qPCR: *CNN3*, *STON2*, *CALN1*, *RUNX2*, *GADD45A*, *CDC45*, *CDC20*, *PLK1*, *AIDA*, *HCK*, *LY86*, *GPR65*, *PIK3CG*, *LILRB2*, *IL7R*, *TCL1A*, *DOCK1*, *HIST1H3G*, *PTPN14*, *CD72* y *NT5E*. Los resultados del GSEA (Gene Set Enrichment Analysis) e IPA (Ingenuity Pathway Analysis) mostraron que el ciclo celular estaba alterado en MO, con una firma de cuatro genes sobreexpresados (*PLK1*, *CDC20*, *CDC45* y *GADD45A*), junto con la baja expresión de *IL7R* y *PIK3CG* que activan importantes vías de señalización para la diferenciación de células B. Un análisis bioinformático comparativo de 15 aspirados de MO y 10 muestras de SP de pacientes Hispánicos diagnosticados con LAL-B colectados por

el Programa TARGET (Therapeutically Applicable Research to Generate Effective Treatments), corroboró los genes observados, excepto por *PI3KCG*. El análisis ARACNE (Algorithm for the Reconstruction of Accurate Cellular Networks) infirió una red que comprendía los genes *AIDA*, *STAP-2*, *FGF19* y *SMIM10LB2*, un lncRNA no reportado en LAL pre-B. En conclusión, los pacientes mexicanos estudiados presentaron alteraciones precursoras en común con los pacientes Hispánicos de la base TARGET pero mutaciones histotipo-específicas que podrían facilitar la estratificación de riesgo y un diagnóstico más preciso. Por otro lado, las proteínas de la maquinaria mitótica y de la diferenciación de células B revelados por este estudio podrían servir como blancos terapéuticos potenciales en LAL pre-B.

## ABSTRACT

Mexico City has one of the highest incidences of B acute lymphoblastic leukemia (B-ALL) globally, with patients showing a poor response to conventional therapy, low survival, and high relapse rates. To gain more insight into the molecular features of B-ALL in Mexican children, we isolated CD10+/CD19+ precursor B lymphoblasts from four bone marrow and nine peripheral blood samples of B-ALL patients using a fluorescence-activated cell sorting (FACS) protocol. The global gene expression profile (BM vs PB) revealed 136 differentially expressed genes; 62 were upregulated (45.6%) and 74 were downregulated (54.4%). Pearson's correlation coefficient was calculated to determine the similarity between pre-B lymphoblast populations. We selected 26 highly significant genes and validated 21 by RT-qPCR (*CNN3*, *STON2*, *CALN1*, *RUNX2*, *GADD45A*, *CDC45*, *CDC20*, *PLK1*, *AIDA*, *HCK*, *LY86*, *GPR65*, *PIK3CG*, *LILRB2*, *IL7R*, *TCL1A*, *DOCK1*, *HIST1H3G*, *PTPN14*, *CD72*, and *NT5E*). The gene set enrichment analysis (GSEA) of the total expression matrix and the ingenuity pathway analysis (IPA) of the 136 differentially expressed genes showed that the cell cycle was altered in the bone marrow with four overexpressed genes (*PLK1*, *CDC20*, *CDC45*, and *GADD45A*) and a low expression of *IL7R* and *PIK3CG*, which are involved in B cell differentiation. A comparative bioinformatics analysis of 15 bone marrow and 10 peripheral blood samples from Hispanic B-ALL patients collected by the TARGET program, corroborated the genes observed, except for *PIK3CG*. The algorithm for the reconstruction of accurate cellular network analysis (ARACNE) inferred a network comprising *AIDA*, *STAP-2*, *FGF19*, and *SMIM10LB2*, an lncRNA not previously



reported in pre-B ALL. We conclude that the Mexican and the Hispanic B-ALL patients studied present common driver alterations but also histotype-specific mutations that could facilitate risk stratification and diagnostic accuracy and serve as potential therapeutic targets.

**Keywords:** ARACNE, FACS, TARGET, pediatric precursor B-ALL, *PIK3CG*, *SMIM10LB2*

## INTRODUCCIÓN

La leucemia es el cáncer infantil más frecuente en el mundo. En el caso de la Leucemia Aguda Linfoblástica (LAL), el resultado del tratamiento ha mejorado de manera consistente durante los últimos 50 años.<sup>1</sup> En las regiones desarrolladas del mundo la supervivencia de 5 años es >90% y la tasa de cura es del 85%. Sin embargo, en países en desarrollo como México, *Jiménez-Hernández et al. (2015)*<sup>2</sup> observaron una baja tasa de supervivencia (64%) y una tasa de recaída de 26.2%; además, el Global Cancer Observatory (GLOBOCAN, IARC, WHO, 2018)<sup>3</sup> reporta que México tiene altas tasas de incidencia estandarizadas por edad por sexo (5.6 para hombres y 4.5 para mujeres por 100,000 habitantes). Aún más, la CDMX tiene una de las incidencias más altas del mundo (5.76 per-100,000), similar a la de los Hispanos que viven en EEUU.<sup>4</sup> Por tanto, además de la clasificación de riesgo en los nuevos casos diagnosticados con LAL usando variables clínicas, citogenéticas, inmunológicas y moleculares, se han conducido varios estudios sobre la Oncología molecular de la LAL y la diversidad genética mestiza en México para estudiar las complejas y heterogéneas características clínicas y biológicas de esta malignidad hematológica en niños mexicanos<sup>5</sup> que también presentan una pobre respuesta a la terapia convencional. Es así que, considerando que la mayoría de los perfiles de expresión LAL-B han sido realizados usando células mononucleares (linfocitos B y T y monocitos) a partir de muestras de pacientes, el objetivo de este estudio fue llevar a cabo el análisis de expresión génica global de linfoblastos precursores B (pre-B) CD10+/CD19+ de muestras de sangre periférica (SP) y aspirados de médula ósea (MO) de pacientes con LAL-B para evitar la

contaminación con otras células mononucleares o no leucémicas. Como *Staal et al.* han reportado, la elección de la técnica y purificación de células influyen la identificación de marcadores de diagnóstico potenciales<sup>6</sup>, así que esta estrategia nos permitió obtener datos más exactos de ambas poblaciones de linfoblastos pre-B leucémicos en el contexto de la enfermedad. Se usaron herramientas bioinformáticas para determinar los procesos celulares y vías de señalización donde los genes diferencialmente expresados (GDEs) estuvieran involucrados. Por último, para comparar y corroborar los hallazgos en el grupo de pacientes mexicanos, y tomando en cuenta que los niños mexicanos tienen un 85% de descendencia Hispana, se efectuó un análisis bioinformático comparativo usando datos de muestras de MO y SP de pacientes Hispanos con B-ALL colectados en la base de datos pública TARGET. Los resultados nos permitieron extraer conclusiones sólidas sobre las similitudes y diferencias entre ambos grupos de pacientes y subraya una vez más la importancia de la vía PI3K/Akt/mTOR pathway en la LAL-B y apunta a proteínas específicas del ciclo celular como blancos terapéuticos para tratar esta malignidad pediátrica.

## **JUSTIFICACIÓN**

Es necesario seguir estudiando la Leucemia Linfoblástica Aguda B (LAL-B) para explicar porqué la Ciudad de México tiene altas tasas de incidencia y los pacientes pediátricos mexicanos presentan una pobre respuesta a la terapia convencional, una tasa de recaída alta (26.2%) y una tasa de supervivencia baja (64%).

## OBJETIVOS

### *General*

Realizar un análisis de expresión global de linfoblastos pre B CD10+/CD19+ de muestras de Sangre Periférica (SP) y Médula Ósea (MO) de pacientes pediátricos mexicanos diagnosticados con LAL-B.

### *Particulares*

1. Implementar un protocolo de FACS (Flourescence Activated Cell Sorting) para purificar linfoblastos pre-B CD10+/CD19+ de muestras de Sangre Periférica (SP) y Médula Ósea (MO) de pacientes pediátricos mexicanos diagnosticados con LAL-B.
2. Realizar un análisis bioinformático: a) de los resultados del análisis de expresión global mediante el software TAC 4.0 (Transcriptome Analysis Console), b) de los genes diferencialmente expresados (GDEs) através del software IPA 3.0 (Ingenuity Pathway Analysis), c) de la matriz de expresión total de TAC mediante un GSEA (Gene Set Enrichment Analysis), d) a partir de una Tabla MS-Excel que sumarice todos los datos, incluyendo los radios fold change Log2 calculados (Muestras MO vs Muestras SP) para cada gen y datos normalizados de fluorescencia, efectuar un análisis ARACNE (Algorithm for the Reconstruction of Accurate Cellular Networks).
3. Corroborar los resultados del análisis de expresión global de linfoblastos pre-B CD10+/CD19+ de muestras de Sangre Periférica (SP) y Médula Ósea (MO) de pacientes pediátricos mexicanos diagnosticados con LAL-B mediante un análisis bioinformático comparativo con muestras de pacientes Hispanos colectadas en la

base de datos TARGET (Therapeutically Applicable Research to Generate Effective Treatments).

## **HIPOTESIS**

1. Las poblaciones de linfoblastos pre-B CD10+/CD19+ de muestras de Sangre Periférica (SP) y Médula Ósea (MO) de pacientes pediátricos mexicanos diagnosticados con LAL-B purificadas por FACS no presentarán contaminación de otras células no leucémicas, lo que permitirá obtener datos más exactos de ambas poblaciones e identificar potenciales marcadores de diagnóstico y tratamiento con mayor certeza.
2. El análisis bioinformático comparativo entre los resultados de los pacientes mexicanos diagnosticados con LAL-B y los pacientes Hispanos con LAL-B de la base de datos TARGET, revelará similitudes y diferencias de expresión génica entre ambos grupos en el contexto de la enfermedad.

## **MATERIAL Y MÉTODOS**

### ***Pacientes***

El presente estudio incluyó 11 pacientes (< 16 años de edad) ingresados en el Servicio de Oncología Pediátrica del Hospital Juárez de México, el Servicio de Hematología del Hospital de Pediatría del CMN sXXI del IMSS y el Servicio de Oncología del Hospital Pediátrico Moctezuma y recién diagnosticados con LAL-B por morfología e inmunotipificación de células de muestras de sangre periférica (SP) y aspirados de médula osea (MO). Fueron estratificados en dos subgrupos:

Riesgo Alto (0-12 meses o >10 años de edad; conteo de leucocitos  $>50 \times 10^9/L$ ) o Riesgo Estandar (1-10 años de edad; conteo de leucocitos  $<50 \times 10^9/L$ ). De manera paralela, se realizó la detección de los re-arreglos cromosómicos *ETV6-RUNX1*, *E2A-PBX1*, *BCR/ABL* y *MLL4/AF4* para cada paciente de acuerdo a *Bekker et al. (2014)*.<sup>7</sup> Se colectaron nueve muestras de SP y cuatro aspirados de MO de los pacientes antes de la terapia de inducción a remisión. El estudio fue aprobado por los Comités de Investigación y Ética (R-2011-785-064) del Instituto Mexicano del Seguro Social, (IMSS). Se obtuvo consentimiento informado escrito y firmado de los padres o guardianes de cada paciente de acuerdo a la Declaración de Helsinki.

### ***Extracción de células mononucleares***

Se colectaron entre 3-5 ml de SP o MO de los pacientes en tubos Vacutainer con EDTA (BD, NJ, USA); inmediatamente después, las muestras fueron separadas usando Histopaque-1077 siguiendo las instrucciones del fabricante (Sigma Aldrich, St. Louis MO, USA). La fase de células mononucleares se colectó cuidadosamente para evitar contaminación con otras interfases (Ver Protocolo 1).

### ***Inmunotipificación y extracción de linfoblastos pre-B CD10+/CD19+ mediante FACS*** (Fluorescence Activated Cell Sorting).

Para la inmunotipificación de los linfoblastos pre-B CD10+/CD19+ se escogieron los siguientes antígenos de superficie: CD45, CD34, CD10 y CD19.

Así, las células mononucleares extraídas por Histopaque-1077 fueron resuspendidas en PBS (1X) y teñidas con anti-CD45 conjugado con PB (Pacific Blue, longitud de onda de emisión 455 nm) (HI30), anti-CD34 conjugado con PE (Ficoeritrina, longitud de onda de emisión 575 nm) (561), anti-CD10 conjugado con FITC (Fluorescein-Isotiocianato, longitud de onda de emisión 520 nm) (HI10a) y anti-CD19 conjugado con APC (Aloficocianina, longitud de onda de emisión 660 nm) (HIB19) (Biolegend, San Diego, CA, USA). Se incubaron por 30 minutos a 4°C y luego se analizaron y purificaron por citometría de flujo activada por fluorescencia (FACS, fluorescence activated cell sorting) en un citómetro de flujo FACS Aria II (BD Biosciences, San Jose, CA, USA). Los linfocitos fueron identificados por sus propiedades *side scatter* y *forward scatter* y se seleccionaron los CD45<sup>+</sup> CD10<sup>+</sup> CD34<sup>+</sup> CD19<sup>+</sup> para obtener finalmente los linfoblastos pre-B CD10<sup>+</sup>/CD19<sup>+</sup>. La pureza fue de 90-95% para todas las muestras y se centrifugaron a 5,000 rpm/ 3 minutos y lavaron con PBS antes de el aislamiento de RNA total (Ver Protocolo 2).

### ***Aislamiento de RNA total***

El RNA total fue extraído por Trizol y resuspendido en agua libre de RNasa (Invitrogen, Carlsbad, CA, USA) de acuerdo a las instrucciones del fabricante. Las muestras con un ratio A260/280 ratio  $\geq 1.8$  se seleccionaron y almacenaron a -70°C hasta su uso (Ver Protocolo 3).

## **Análisis de Microarreglos**

La integridad del RNA se evaluó usando el Bioanalizador Agilent 2100 (Agilent Technologies, Santa Clara, CA, USA). Sólo las muestras con un RIN (RNA integrity number) >7.5 fueron incluidas. Los microarreglos HuGene-1\_0-st-v1 fueron empleados para comparar la expresión génica (Muestras MO vs Muestras SP) de los linfoblastos pre-B CD10+/CD19+. El análisis de microarreglos se realizó para cada muestra de MO o SP siguiendo los protocolos del fabricante (Affymetrix Inc., Santa Clara, CA, USA) (Ver Protocolo 4).

## **Análisis bioinformático**

**TAC.** Los datos de microarreglos fueron analizados con el software TAC 4.0 (Transcriptome Analysis Console) de Affimetrix (Fold Change  $\geq 1.7$  o  $\leq -1.7$ , p-value  $\leq 0.05$ ), el protocolo *Limma* y parámetros por default.

**IPA.** Los datos de los genes diferencialmente expresados (GDEs) fueron analizados através del software IPA 3.0 (Ingenuity Pathway Analysis) (QIAGEN Inc., <https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis>).

**GSEA.** Adicionalmente, la matriz de expresión total de TAC se usó para conducir un GSEA (Gene Set Enrichment Analysis) con las herramientas gratuitas *on line* del Instituto Broad (<http://software.broadinstitute.org/gsea/index.jsp>) empleando gene-set como tipo de permutación y parámetros por default. Los GDEs se usaron para elaborar un heatmap con la herramienta *on line HeatMapper* utilizando el Clustering method centroid linkage and Distance Measurement con el Coeficiente de Correlación de Pearson.<sup>8</sup>



**ARACNE.** Se preparó una Tabla MS-Excel Table sumando todos los datos incluyendo los ratios fold change Log2 calculados (Muestras MO vs Muestras SP) para cada gen y datos normalizados de fluorescencia. Los datos normalizados de fluorescencia por muestra se procesaron mediante ARACNE (Algorithm for the Reconstruction of Accurate Cellular Networks) que consiste en el cálculo de la información mutua (MI, mutual information) entre pares de genes para identificar posibles interacciones.<sup>9</sup> Se usó el plugin ARACNE disponible para Cytoscape v3.1.1.<sup>10</sup> El cálculo de red (network) se efectuó usando el Aracne Complete Mode, con el algoritmo de MI Variable Bandwidth, tolerancia de DPI (Data Processing Inequality) de 0 y 6 pasos de MI. Se probaron diferentes valores umbral MI de 0.6 hasta 0.9

### ***Análisis de Tiempo Real***

Primero se efectuó la transcripción reversa de 100 a 150 nanogramos de RNA total de cada muestra de SP y MO en una reacción de 20 uL siguiendo las especificaciones del fabricante (Invitrogen, Carlsbad, CA, USA). Después, se efectuó la expresión relativa de 26 GDEs mediante PCR cuantitativa en Tiempo Real (RT-qPCR), usando un 7300 Real Time PCR System, placas de reacción óptica de 96 pozos (Applied Biosystems, CA, USA) y primers de oligonucleotidos validados de los genes: *AIDA*, *CALN1*, *CDC20*, *CDC45*, *CNN3*, *GADD45A*, *GPR65*, *HCK*, *IL7R*, *LILRB2*, *LY86*, *PIK3CG*, *PLK1*, *RUNX2*, *STON2*, *TCL1A*, *CD72*, *DOCK1*, *HIST1H3G*, *NT5E*, *PTPN14*, *CCNB2*, *CD86*, *HDAC9*, *HIST1H2B1* y *HIST1H3F* (Prime Time® IDT, San Diego, CA, USA). La amplificación de cada

templado se efectuó por triplicado en una corrida de PCR. Los datos se analizaron usando el método Livak: Cantidad de blanco =  $2^{-\Delta\Delta CT}$ . Se realizó la cuantificación de RNAm de  $\beta 2$ -microglobulina como control endógeno (Ver Protocolos 5 y 6).

***Análisis bioinformático comparativo con muestras de pacientes Hispanos disgnosticados con LAL-B de la base de datos pública TARGET (Therapeutically Applicable Research to Generate Effective Treatments).***

Para corroborar los resultados obtenidos por el perfil de expresión global por microarreglos de los pacientes mexicanos con LAL-B se efectuó un análisis comparativo usando datos de una base de datos pública. Usando la plataforma cBio-Portal<sup>11-13</sup>, se descargaron datos de secuencia (RNAseq-FPKMs) de pacientes pediátricos con LAL-B colectados como parte del Programa TARGET de la Office of Cancer Genomics (OCG) del National Cancer Institute (NIH)<sup>14-15</sup>. Los datos se seleccionaron de acuerdo a los siguientes criterios de inclusión: pacientes menores de 16 años, Hispanos, célula de origen: precursor B (pre-B), muestras de MO o SP colectadas al momento del diagnóstico y previo a la terapia de inducción a la remisión; sólo un total de 25 muestras cumplieron los criterios de inclusión (15 MOs y 10 SPs). Los datos se compararon (muestras MO vs muestras SP) a través de un análisis de expresión diferencial (DEA, por sus siglas en inglés) usando el IDEA mex. Las herramientas y heatmaps iDEP.92, análisis de enriquecimiento y análisis de vías de señalización se realizaron usando las herramientas iDEP.92, ShinyGO v0.66, Pathview y GSEA. El valor de significancia

fue  $p \leq 0.05$ ; el resto de los valores se usó por default. Las bases de datos KEGG y Gene Ontology (GO) se utilizaron para todos los análisis<sup>16-20</sup>.

## **RESULTADOS**

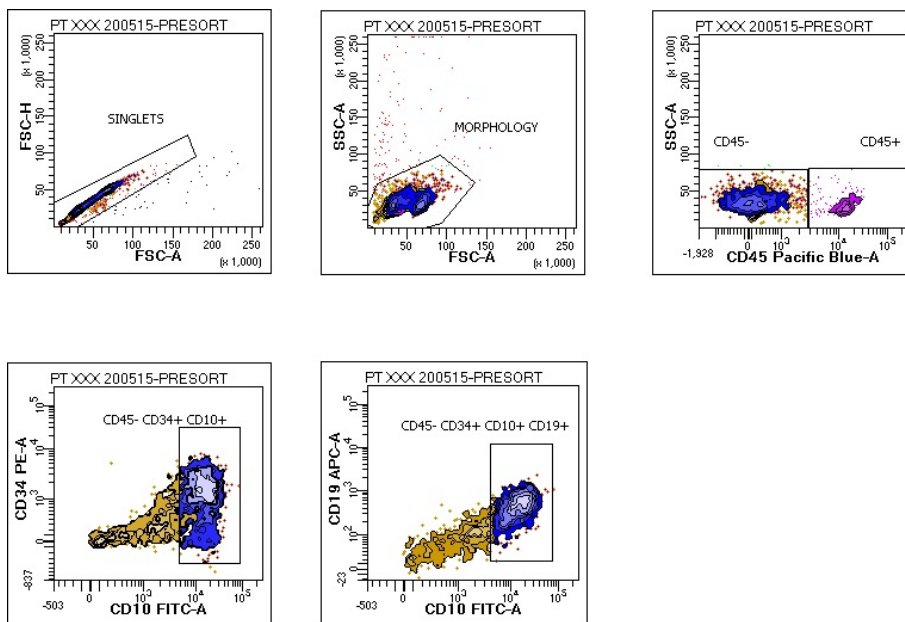
Se implementó un protocolo FACS para purificar linfoblastos pre-B CD10+/CD19+ a partir de MO y SP de pacientes pediátricos con LAL pre-B. Sorprendentemente, las concentraciones de RNA total de menos de  $4 \times 10^6$  linfoblastos pre-B CD10+/CD19+ no pudieron ser detectadas por el espectrofotómetro NanoDrop 2000 (Thermo Scientific., MA, USA) ni por el Bioanalizador para evaluar la integridad del RNA. Consecuentemente, sólo aquellas muestras de MO y SP con al menos  $4 \times 10^6$  o más linfoblastos pre-B CD10+/ CD19+ purificados por FACS fueron seleccionadas para el aislamiento de RNA total en altas concentraciones, pureza e integridad para el análisis de expresión génica global y validación de los genes significativos por RT-qPCR. Por tanto, después de múltiples ensayos, el número final de muestras fue de 13: nueve muestras de SP y cuatro aspirados de MO provenientes de 11 pacientes recién diagnosticados con LAL pre-B (Tabla 1, Figura 1).

\* LAL-B: Leucemia Aguda Linfoblástica B; MO: Médula ósea; SP: Sangre periférica; CCB: Conteo de células blancas; Dx: diagnóstico; NCI: National Cancer Institute; RA: Riesgo alto; RS: Riesgo standard; ND: No detectado. Los rearrreglos cromosómicos fueron detectados como se indica en Material y Métodos.

**Tabla 1. Datos y variables clínicas de pacientes recién diagnosticados con LAL-B incluidos en este estudio**

Muestra ID	Sitio	Paciente ID	Sexo	Edad al Dx (meses)	CCB al Dx (x109)/ L	Grupo de riesgo	Rearreglo cromosómico <i>ETV6-RUNX1</i>	Rearreglo cromosómico <i>E2A-PBX1</i>	Rearreglo cromosómico <i>BCR/ABL</i>	Rearreglo cromosómico <i>MLL4/AF4</i>	Recaída temprana	Muerte
1	SP	1	F	123	92.3	RA	ND	ND	ND	ND	SÍ	SÍ
2	SP	2	M	60	30	RS	ND	ND	ND	ND	NO	NO
3	SP	3	M	60	40	RS	ND	ND	ND	ND	NO	NO
4	SP	4	M	47	529	RA	ND	ND	ND	ND	NO	NO
5	SP	5	M	84	53.1	RA	ND	ND	ND	ND	NO	NO
6	SP	6*	M	187	10.24	RA	ND	ND	ND	ND	NO	NO
7	SP	7**	M	64	19.28	RS	ND	ND	ND	ND	NO	NO
8	SP	8	M	48	35.17	RS	ND	ND	ND	ND	NO	NO
9	SP	9	M	132	123.4	RA	ND	ND	ND	ND	NO	NO
					<b>BLASTOS al Dx (%)</b>							
10	MO	6*	M	187	98	RA	ND	ND	ND	ND	NO	NO
11	MO	10	F	156	99	RA	ND	ND	ND	ND	NO	NO
12	MO	11	F	168	100	RA	<b>DETECTADO</b>	ND	ND	ND	SÍ	SÍ
13	MO	7**	M	64	95	RS	ND	ND	ND	ND	NO	NO

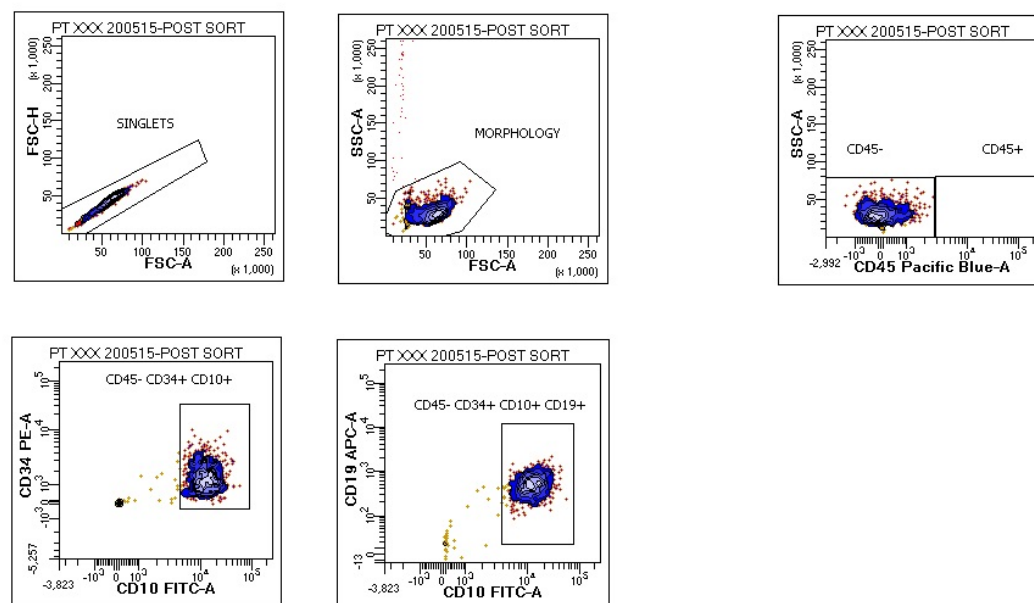
A



Tube: PRESORT

Population	#Events	%Parent	%Total
All Events	2,000	###	100.0
SINGLETs	1,957	97.9	97.9
MORPHOLOGY	1,823	93.2	91.1
CD45-	1,585	86.9	79.2
CD45- CD34+ CD10+	957	60.4	47.8
CD45- CD34+ CD10+ CD19+	956	99.9	47.8
CD45+	233	12.8	11.6

B



Tube: POST SORT

Population	#Events	%Parent	%Total
All Events	1,000	###	100.0
SINGLETs	1,000	100.0	100.0
MORPHOLOGY	959	95.9	95.9
CD45-	958	99.9	95.8
CD45- CD34+ CD10+	917	95.7	91.7
CD45- CD34+ CD10+ CD19+	917	100.0	91.7
CD45+	0	0.0	0.0

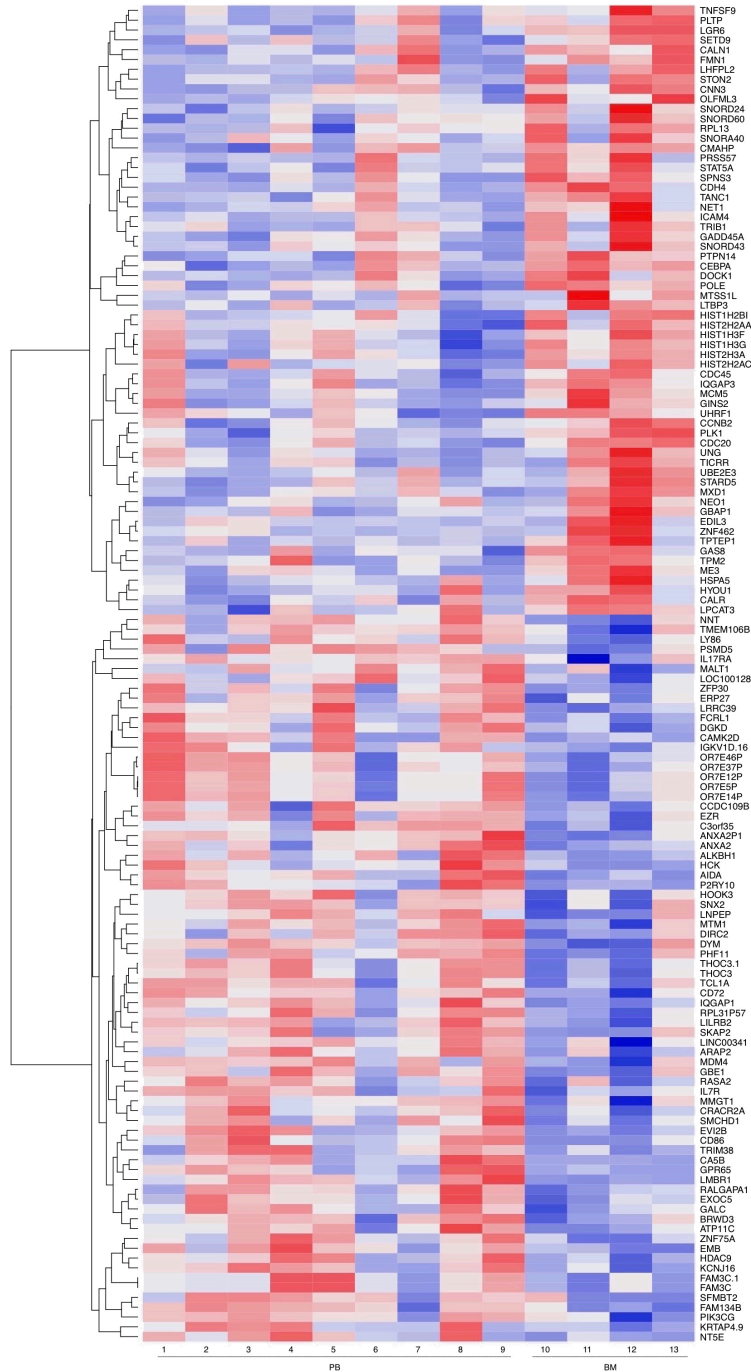
**Figura 1. Inmunotipificación.** Gráficas de dispersión de FACS de células de fase blanca de sangre periférica (SP) o médula ósea (MO) de pacientes diagnosticados con LAL-B y teñidas con anticuerpos CD45-PB, CD34-PE, CD10-FITC y CD19-APC. **A)** Gráfica de dispersión *pre-sort* que muestra a CD45 como anticuerpo ancla y a CD34 para “gating” de leucocitos. **B)** Gráfica de dispersión *post-sort* mostrando los linfoblastos pre-B CD10+/CD19+.

\* PB (Pacific Blue); PE (Ficoeritrina), FITC (Fluorescein-Isotiocianato), APC (Aloficocianina).

El perfil de expresión global de linfoblastos pre-B CD10+/ CD19+ purificados por FACS a partir de nueve muestras de SP y cuatro aspirados de MO se llevó a cabo usando el chip HuGene-1\_0-st-v1 de Affymetrix. El análisis jerárquico de clusters (hierarchical clustering) de los datos de todos los pacientes reveló 136 genes diferencialmente expresados (GDEs): 62 sobrepresados (45.6%) y 74 subexpresados (54.4%); el coeficiente de correlación de Pearson se calculó para encontrar la asociación entre las muestras de SP y MO y entre los GDEs; el heatmap resultante, con la muestras de SP a la izquierda y las muestras de MO a la derecha, mostró diferencias entre ambas poblaciones de linfoblastos pre-B CD10+/ CD19+ (Figura 2).



Heatmap Bone Marrow vs Peripheral Blood



**Figura 2. Microarreglos de expresión génica global.** Heatmap de linfoblastos pre-B CD10+/CD19+ de muestras de sangre periférica (SP; n=9) y médula ósea (MO; n=4) de pacientes pediátricos mexicanos recién diagnosticados con LAL-B. Los datos de 136 genes diferencialmente expresados (GDEs): 62 sobreexpresados (rojo) y 74 subexpresados (azul) se usaron para crear el heatmap con el programa *HeatMapper* mediante clustering method centroid linkage and distance measurement con coeficiente de correlación de Pearson.

Algunos de los genes sobreexpresados (mitad superior de la imagen) están involucrados en el ciclo celular y en la remodelación de cromatina y se expresaron en linfoblastos pre-B CD10+/ CD19+ de MO (rojo) pero no de SP (azul); en contraste, los genes subexpresados (mitad inferior de la imagen), algunos de los cuales están involucrados en diferenciación de las células B, no se expresaron en linfoblastos pre-B CD10+/ CD19+ de MO (azul). Es decir, los linfoblastos pre-B CD10+/ CD19+ de MO proliferan y no se diferencian y los linfoblastos pre-B CD10+/ CD19+ de SP no proliferan fuera como lo hacen dentro del nicho hematopoyético. En concordancia con esto, y como consecuencia del procedimiento metodológico que requería concentraciones de RNA que pudieran ser detectadas por los instrumentos, todas las muestras de MO mostraron altos porcentajes de blastos (95-100) y cuatro muestras de SP (IDs 1, 4, 5 and 9) mostraron muy altos conteos de linfocitos ( $>50 \times 10^9/L$ ) lo que se corresponde con una alta proliferación en la MO y subsecuente liberación en el torrente sanguíneo. Es notable que el rearrreglo cromosómico ETV6-RUNX1 fue detectado sólo en un paciente femenino (ID 12) que recayó y eventualmente falleció. El resto de las muestras no mostró translocaciones, sin embargo, un paciente femenino (ID 1) también recayó y murió. De manera interesante, un paciente masculino (ID 4) con hiperleucocitosis ( $529 \times 10^9/L$ ) no recayó y está siendo monitoreado de manera regular. Los resultados de GSEA e IPA ( $p\text{-value} \leq 0.05$ ; Ver Fig. S1, Fig. S2 y Fig. S3) hicieron posible dilucidar aún más las funciones biológicas de la célula y/o vías involucradas y permitieron la selección de 26 de los principales genes significativos para validar mediante RT-qPCR. Se observaron resultados similares

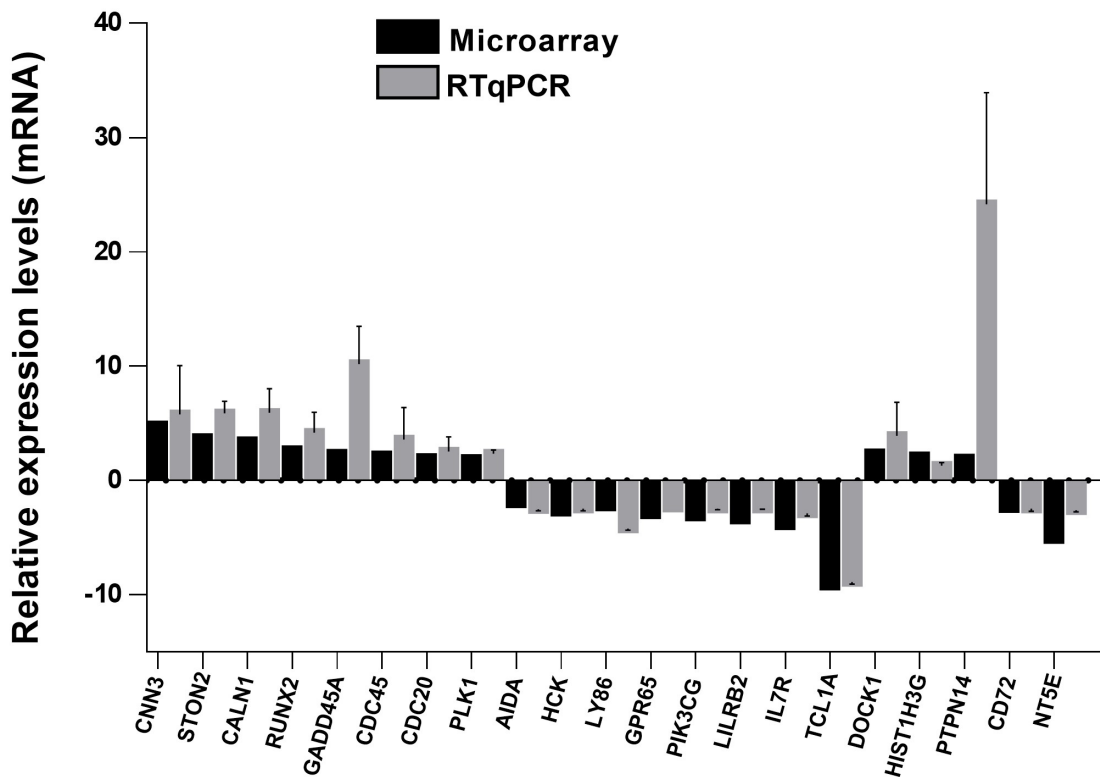


en los niveles de expresión de 21 genes cuyos datos de microarreglos se compararon con los que se obtuvieron por RT-qPCR (Tabla 2 y Figura 3).

**Tabla 2. Genes diferencialmente expresados (GDEs).** Principales genes significativos sobreexpresados y subexpresados en linfoblastos pre-B CD10+/CD19+ purificados por FACS de muestras de SP y MO de pacientes con LAL-B y validados por RT-qPCR.

Símbolo del gen	Descripción	Microarreglos		RT-qPCR	
		Fold change (>1.7 or <1.7)	p ≤ 0.05	Fold change	p ≤ 0.05
<b>CNN3</b>	Calponin 3, acidic	4.8	0.0248	5.79	0.0354
<b>STON2</b>	Stonin 2	3.72	0.0275	5.85	0.0015
<b>CALN1</b>	Calneuron 1	3.43	0.0349	6.57	0.0074
<b>RUNX2</b>	Runt-related transcription factor 2	2.65	0.048	4.16	0.0499
<b>DOCK1</b>	Dedicator of Cytokinesis 1	2.41	0.03	3.9	0.044
<b>GADD45A</b>	Growth arrest and DNA-damage-inducible, alpha	2.36	0.0361	10.17	0.0015
<b>CDC45</b>	Cell Division Cycle 45	2.23	0.05	3.6	0.0284
<b>HIST1H3G</b>	Histone cluster 1 H3 Family Member G	2.1	0.05	1.27	0.001
<b>CDC20</b>	Cell division cycle 20	1.98	0.0048	2.53	0.0283
<b>PLK1</b>	Polo-like kinase 1	1.9	0.0017	2.35	0.032
<b>PTPN14</b>	ProteinTyrosine Phosphatase, Non-Receptor Type 14	1.94	0.024	24.17	0.036
<b>AIDA</b>	Axin interactor, dorsalization associated	-2	0.0093	-2.49	0.0286
<b>HCK</b>	HCK proto-oncogene, Src family tyrosine kinase	-2.06	0.0171	-2.46	0.0444
<b>LY86</b>	Lymphocyte antigen 86	-2.28	0.031	-4.22	0.0333
<b>CD72</b>	Cluster of Differentiation 72	-2.42	0.012	-2.48	0.0452
<b>GPR65</b>	G protein-coupled receptor 65	-2.95	0.044	-2.37	0.001
<b>PIK3CG</b>	Phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit gamma	-3.16	0.0284	-2.44	0.0378
<b>LILRB2</b>	Leukocyte immunoglobulin-like receptor, subfamily B, member 2	-3.41	0.0157	-2.47	0.024
<b>IL7R</b>	Interleukin 7 Receptor	-3.95	0.05	-2.86	0.0284
<b>NT5E</b>	5'-Nucleotidase Ecto	-5.13	0.031	-2.58	0.0303
<b>TCL1A</b>	T-cell leukemia/lymphoma 1A	-9.22	0.0429	-8.9	0.0284

\* MO: Médula ósea; SP: Sangre periférica; LAL-B: Leucemia Aguda Linfoblástica B.



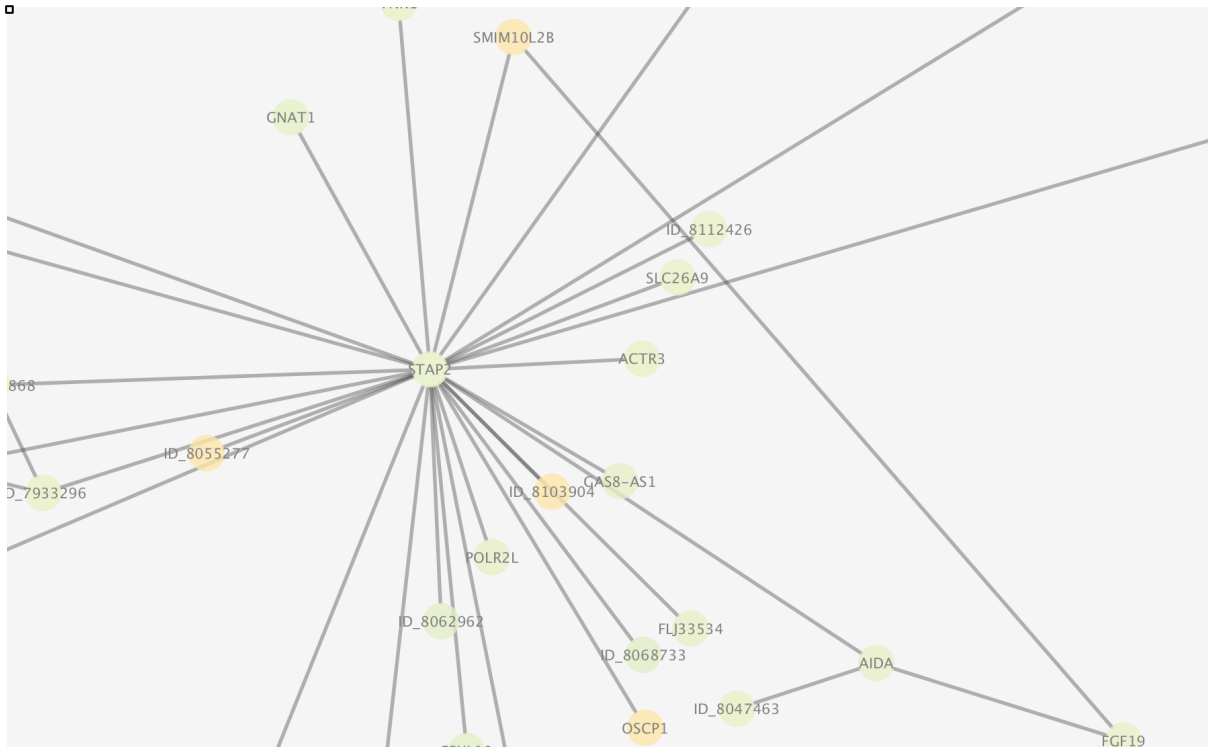
**Figura 3. Validación de datos de microarreglos por RT-qPCR.** Comparación de los niveles de expresión de los genes seleccionados determinados por el análisis de microarreglos y PCR tiempo real. La expresión relativa se determinó por el método  $2^{-\Delta\Delta CT}$  y el promedio de los valores de las muestras de sangre periférica (SP) se usó como calibrador. Los datos de RT-qPCR están expresados como la media/DE de tres análisis independientes. Las barras negras representan los valores de microarreglos y las barras grises los valores de RT-qPCR.

\* DE: Desviación estándar.

En concordancia con *Knaack et al. (2014)*<sup>21</sup> que realizaron un análisis de red reguladora modular *pan-cancer* de seis cánceres humanos y concluyeron que éstos comparten una red reguladora común, se encontró que los GDEs validados

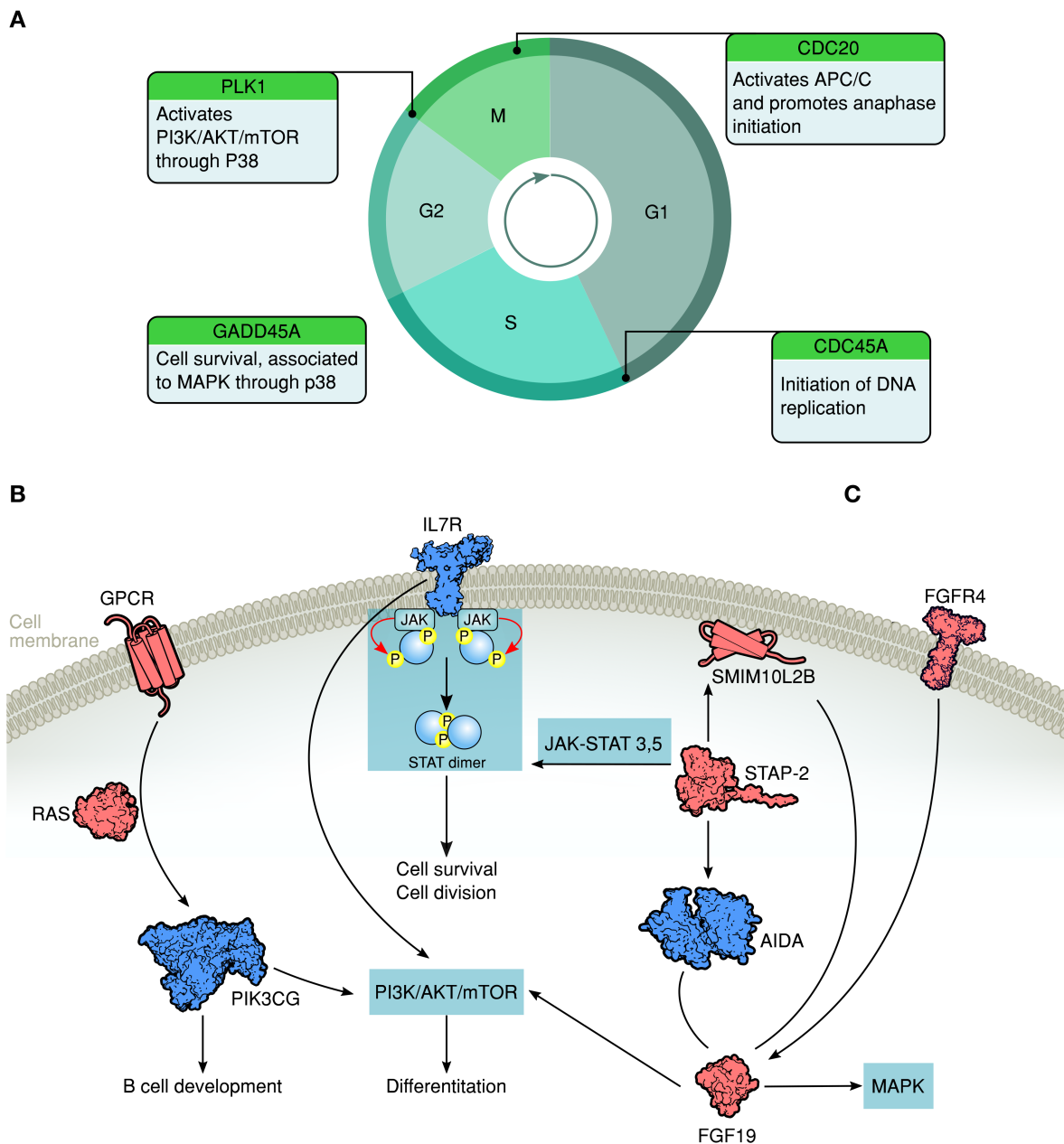
en este estudio están involucrados en ciclo celular (*CDC20*, *PLK1*, *GADD45A*, *CDC45*), respuesta inmune (*TCL1A*, *GPR65*, *LILRB2*, *LY86*, *CD72*, *NT5E*), remodelación de cromatina (*HIST1H3G*), tráfico de vesículas (*STON2*, *CNN3*) y son parte de o codifican para cinasas importantes *corriente arriba* o *corriente debajo* de importantes vías de señalización de la diferenciación de la célula B (*RUNX2*, *PIK3CG*, *IL7R*, *HCK*). Además, es de notar que la mayoría de estos genes se encontró que están principalmente asociados a las vías de señalización PI3K/Akt/mTOR, MAPK y JAK/STAT.

Por otro lado, el análisis ARACNE (valores umbral MI de 0.6 a 0.9) infirió una red con *AIDA* (Figura 4), uno de los genes validados, y tres genes más: *STAP2*, *FGF19* y *SMIM10L2B*, un RNA largo no codificante (lncRNA). Es de mencionarse el hecho de que *AIDA* mostró resultados similares a valores umbral MI 0.7 y 0.8. Adicionalmente, el análisis ARACNE (valor umbral MI 0.7) infirió asociaciones para dos de los genes del ciclo celular observados en este estudio: *PLK1* se asoció a *OR11H7* (Fig. S4), un receptor olfatorio, y *CDC45* se asoció a *KCNQ1DN* (Fig. S5), un gen afiliado también a la clase lncRNA.



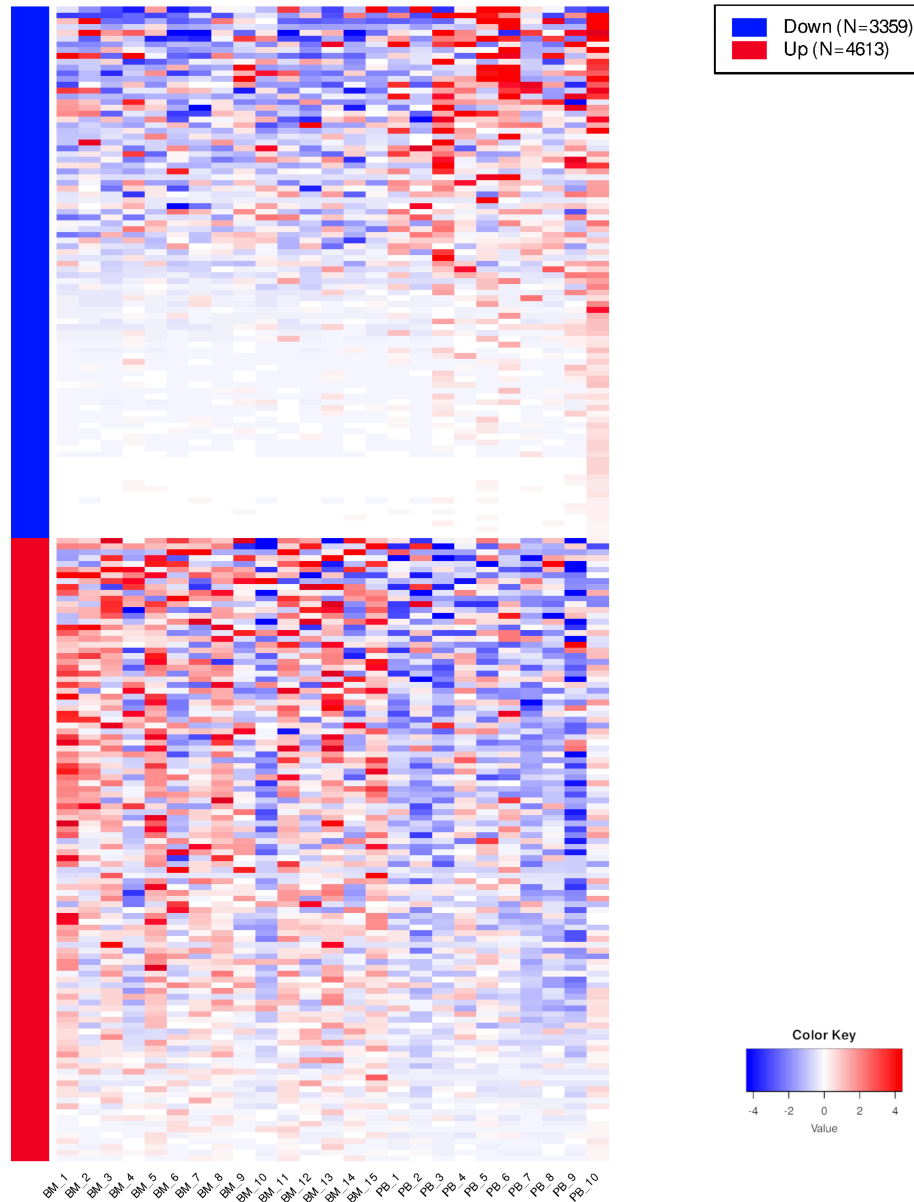
**Figura 4. Red ARACNE (MI 0.70) para el gen *AIDA*: *STAP2*, *FGF19* y *SMIM10L2B*.** Los colores de los nodos varían de azul/ verde (fold changes log2 positivos; esto es, niveles más elevados de estos genes en las muestras de MO) a naranja/rojo (fold changes log2 negativos; esto es, niveles más elevados de estos genes en las muestras de SP).

La Figura 5 resume las observaciones más importantes de este perfil de expresión génica global de linfoblastos pre-B CD10+/CD19+ de pacientes pediátricos diagnosticados con LAL-B.

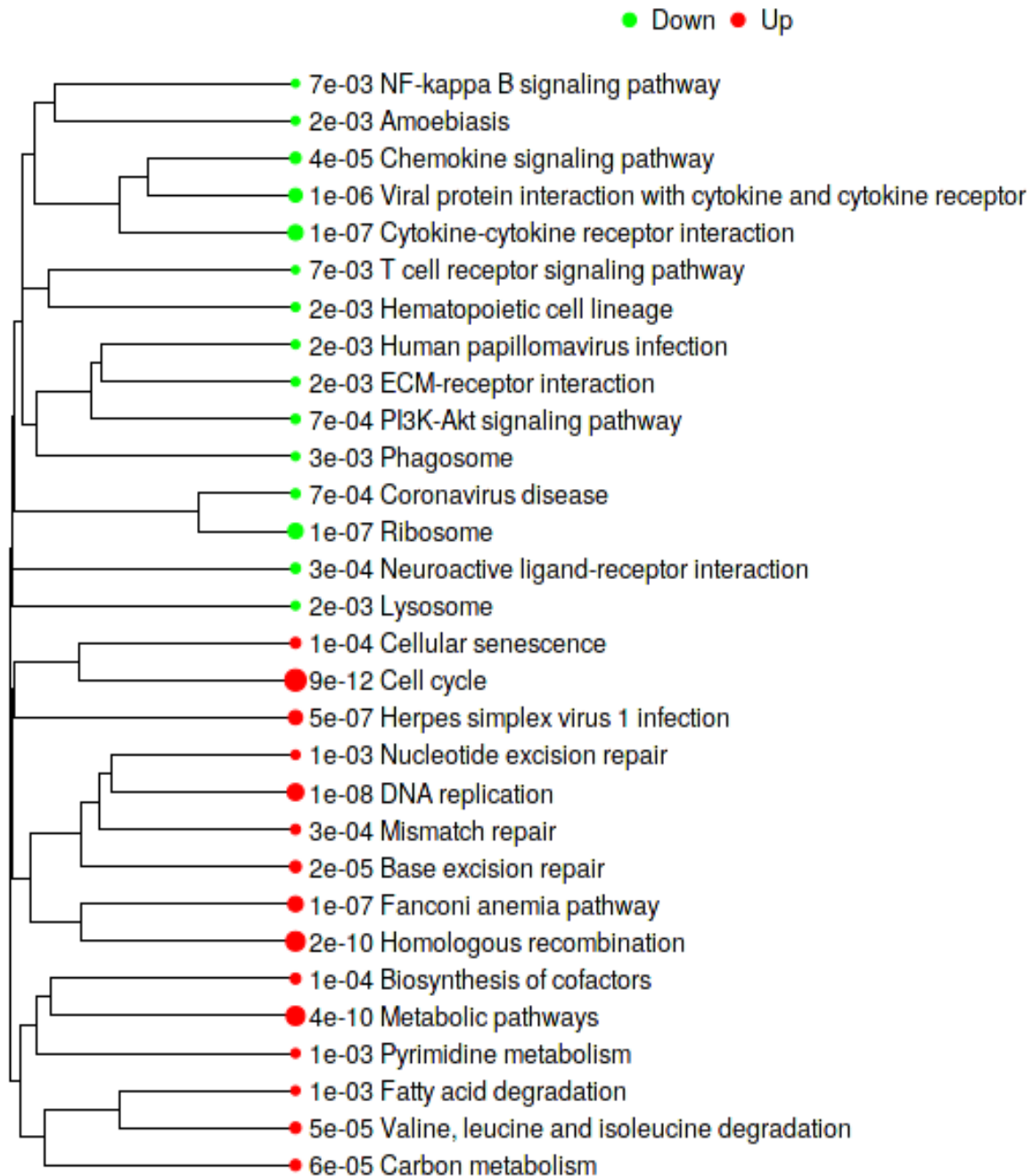


**Figura 5. Firmas génicas. A)** Genes de ciclo celular sobrepresados (*PLK1*, *CDC20*, *GADD45A* y *CDC45*) que aumentan la proliferación celular en linfoblastos pre-B  $CD10^+/CD19^+$  de las muestras de MO pero no las de SP. **B)** Genes subexpresados (*IL7R* y *PIK3CG*, en azul) que inhiben la diferenciación o desarrollo en linfoblastos pre-B  $CD10^+/CD19^+$  de muestras de MO. **C)** Red inferida por el análisis ARACNE para el gen *AIDA*, subexpresado en este estudio. Es de notar *SMIM10L2B*, un lncRNA no reportado antes en LAL-B.

El análisis bioinformático comparativo de 25 muestras (15 MO vs 10 SP) de pacientes pediátricos Hispánicos diagnosticados con LAL-B colectados en la base de datos TARGET, corroboró los hallazgos de este perfil de expresión por microarreglos de linfoblastos pre-B CD10+/CD19+ de pacientes pediátricos mexicanos. Los heatmaps generados mediante las herramientas iDEP e IDEAMEX para el análisis de expresión diferencial (DEA, por sus siglas en inglés) (MO vs SP) de datos de RNA seq (FPKMs), mostró 4,316 genes sobreexpresados y 3,359 subexpresados (Figura 6A), donde el cluster de vías de señalización indicó que los primeros estaban principalmente en el ciclo celular ( $9e-12$ ) y los segundos en la vía PI3K-Akt pathway ( $7e-04$ ) (Figura 6B).



**Figura 6A. Análisis de Expresión Diferencial (AED).** El programa IDEAMEX se usó para realizar un AED (MO vs SP) de datos de RNAseq (FPKMs) de muestras de pacientes Hispanos con LAL-B colectados en la base de datos TARGET. El heatmap resultante mostró 4,316 genes sobreexpresados (rojo) y 3,359 subexpresados (azul). Muestras de MO a la izquierda y muestras de SP a la derecha.

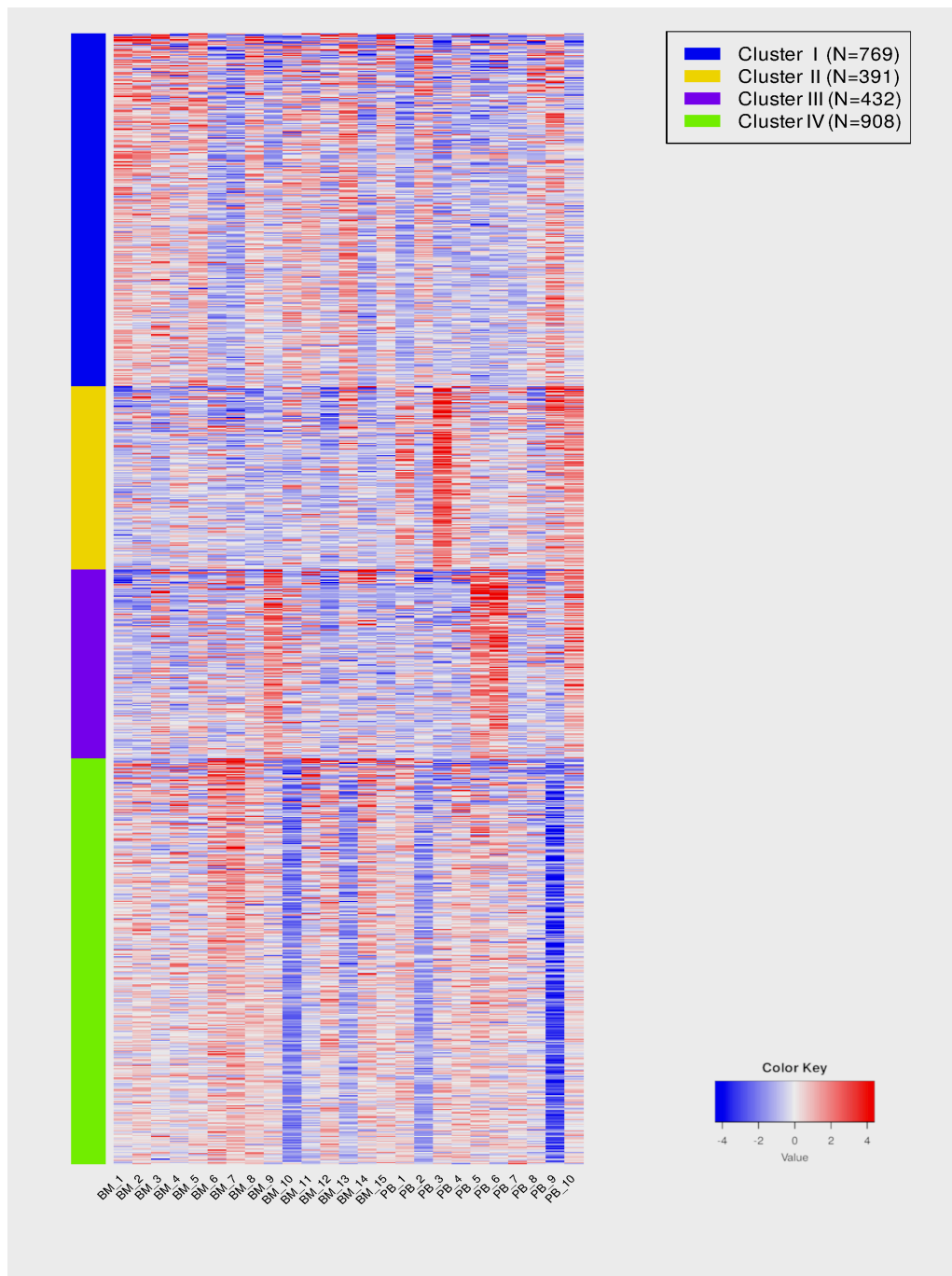


**Figura 6B.** Cluster de vías de señalización mostrando sobreexpresión en los genes del ciclo celular (9e-12, en rojo) y subexpresión en los genes de la vía PI3K-Akt (7e-04, en verde).

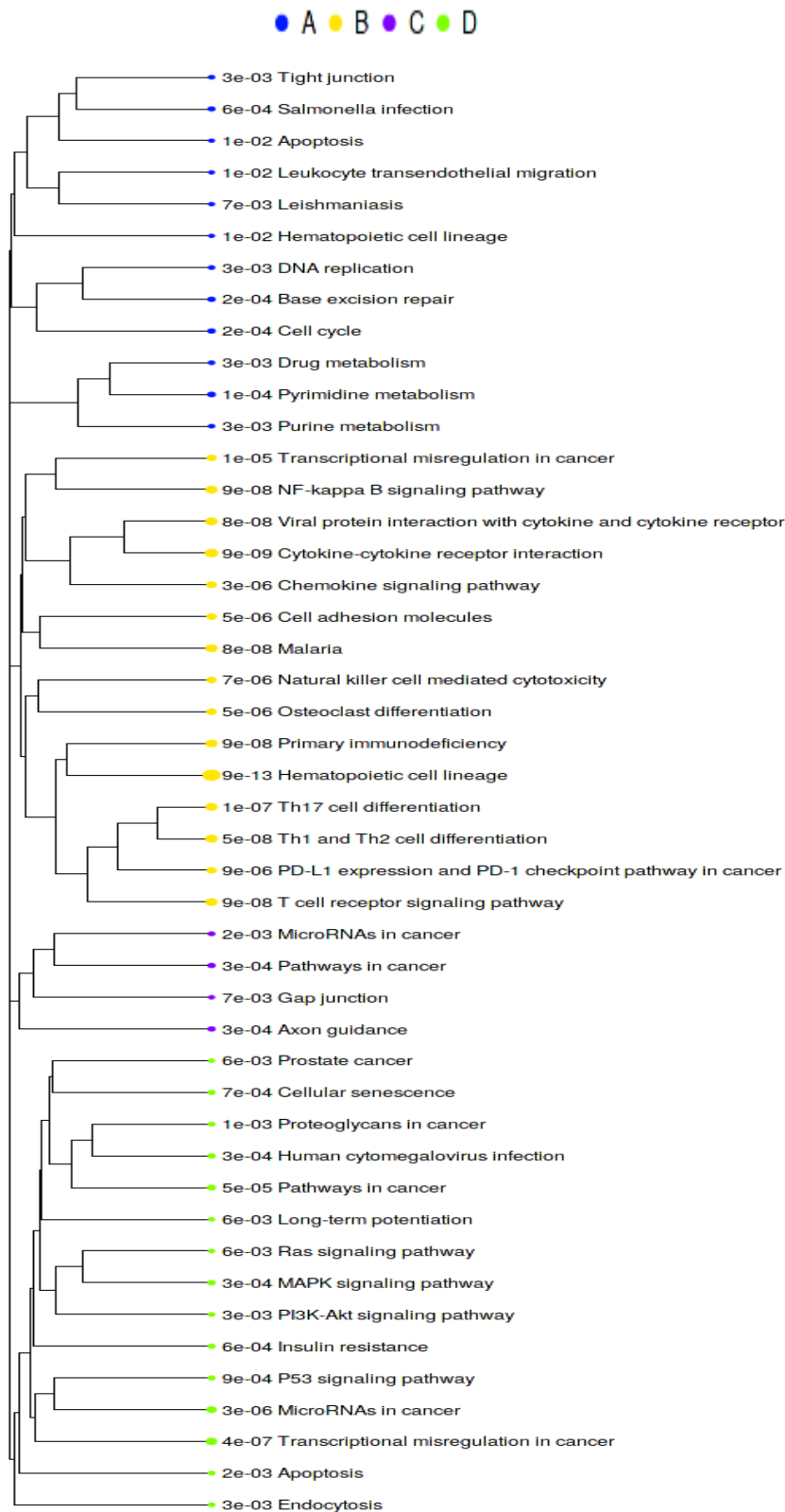


Con la herramienta iDEP.92 se generó un heatmap de cuatro clusters de genes (Figure 7A) y cluster de redes (network clusters) que mostró el ciclo celular ( $2e-04$ ) en el cluster I y las vías de señalización Ras 8 ( $6e-03$ ), MAPK ( $3e-04$ ) y PI3K-Akt ( $3e-03$ ) en el cluster IV (Figure 7B).

A

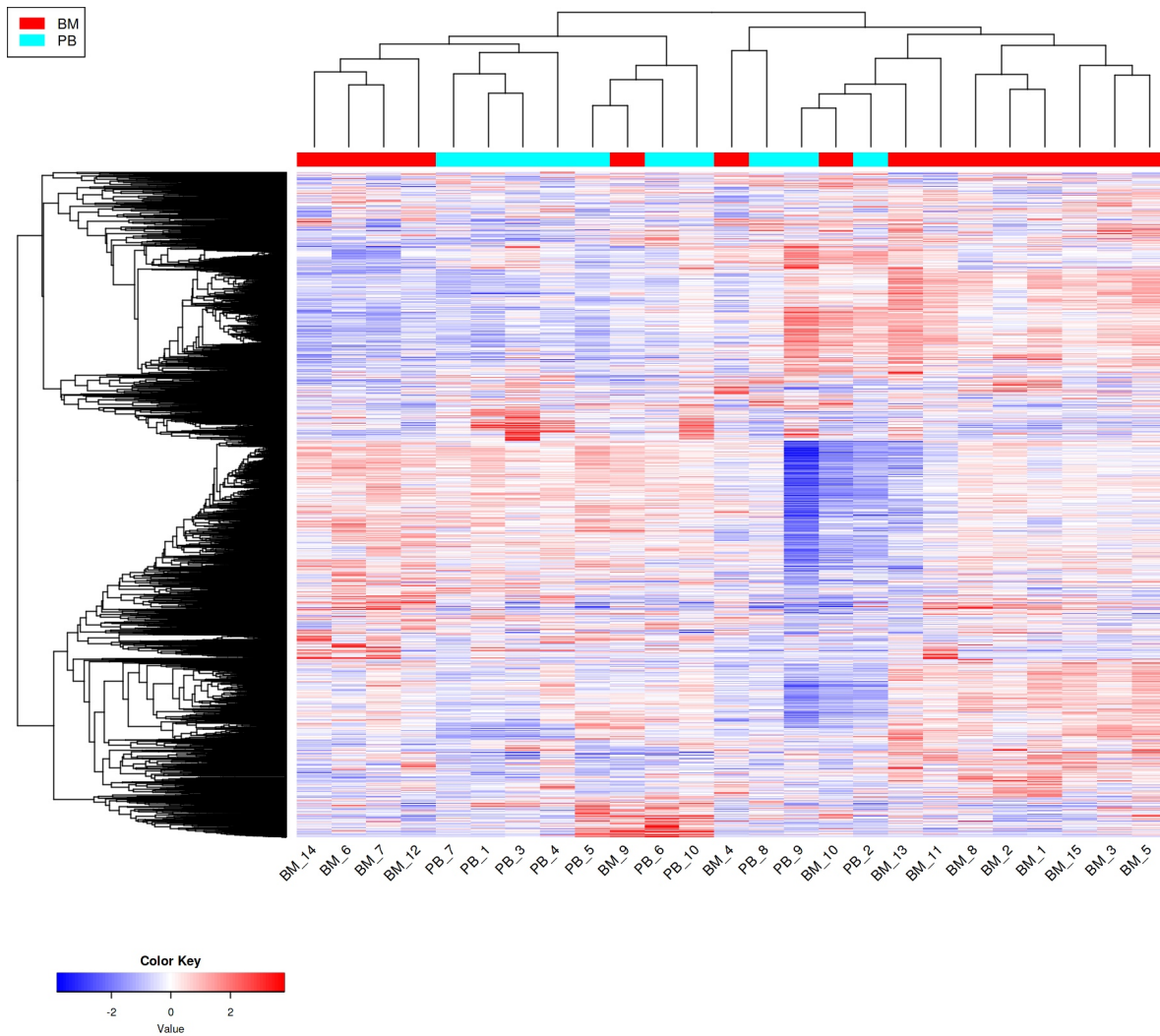


**Figura 7A. Análisis de enriquecimiento.** Heatmap generado a través del programa iDEP.92 mostrando los clusters de genes I, II, III y IV. Muestras de MO a la izquierda y muestras de SP a la derecha.



**Figura 7B. Clusters de redes (networks)** mostrando vías significativas para este estudio: I-A) Ciclo celular y IV-D) Vías de señalización Ras 8, MAPK y PI3k-Akt.

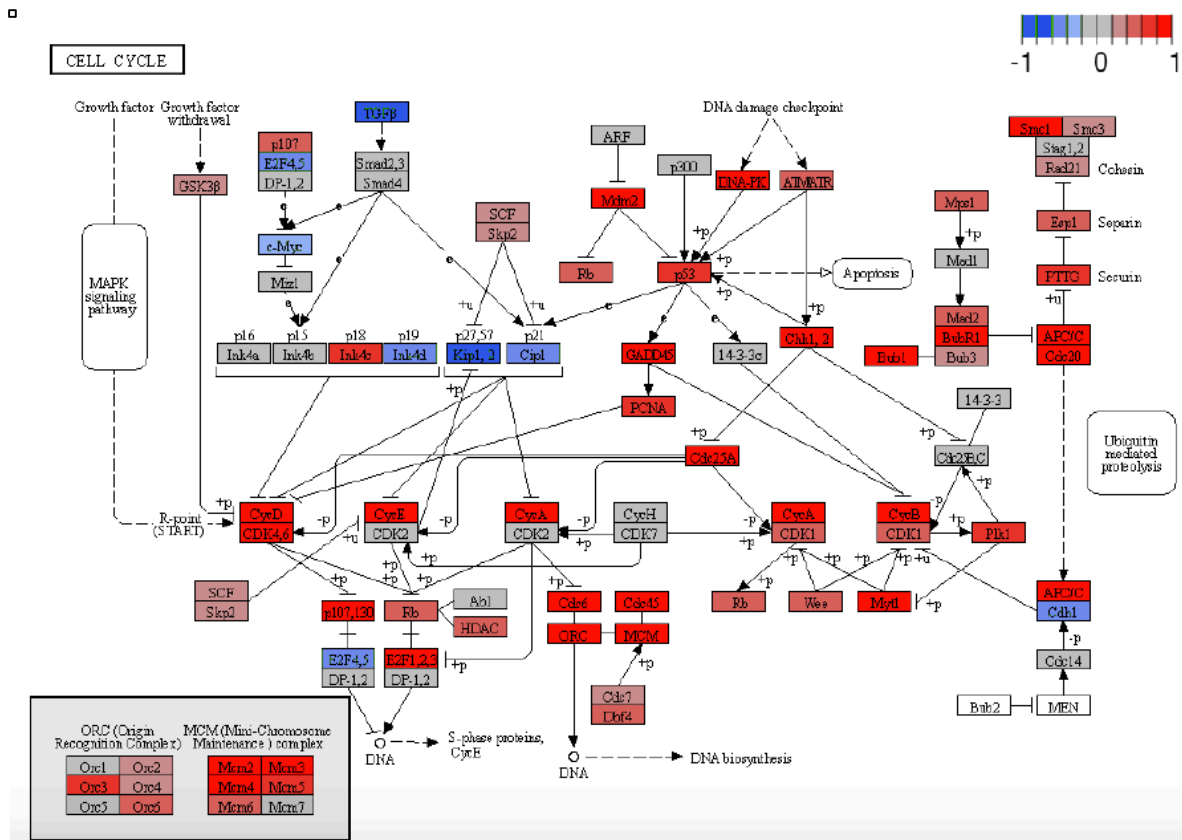
La Figura 8 muestra el heatmap de correlación de los datos de RNA-seq de las muestras de MO vs SP de pacientes Hispánicos de la base de datos TARGET.



**Figura 8.** Heatmap de correlación de datos de RNA seq de las muestras de MO vs. SP de los pacientes Hispánicos con LAL-B de la base TARGET.

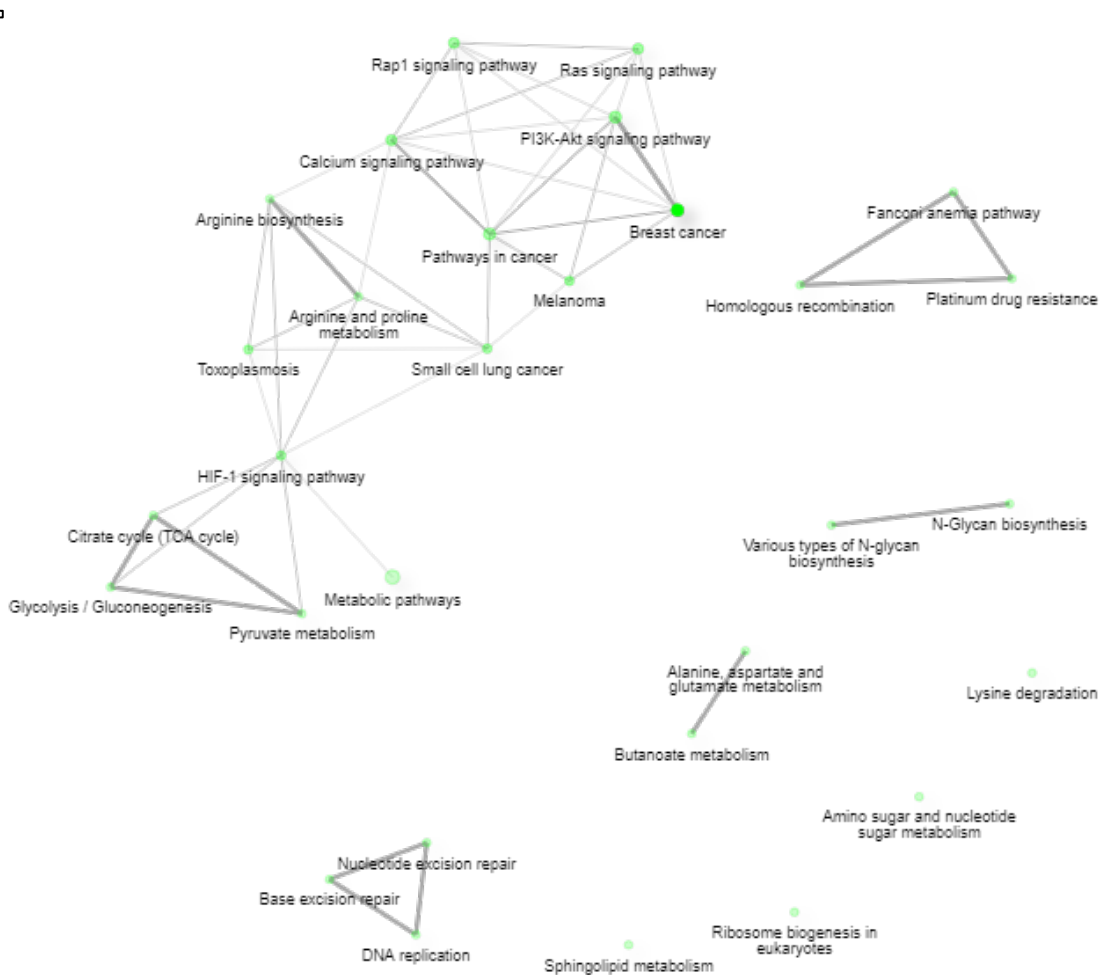
Usando las herramientas iDEP, ShinyGO, Pahtview y GSEA se realizó un análisis de enriquecimiento y de vías de señalización. La Figura 9 muestra la vía del ciclo celular donde la firma de genes sobreexpresados observada en el análisis de

expresión de microarreglos de pacientes mexicanos con LAL-B y validada por RT-qPCR (*CDC20*, *CDC45*, *PLK1* y *GADD45A*), también está presente con niveles altos en las muestras de pacientes Hispánicos con LAL-B de la base TARGET.



**Fig. 9** Vía del ciclo celular mostrando que la firma de genes sobreexpresados observados en el análisis de expresión de microarreglos y validada por RT-qPCR (*CDC20*, *CDC45*, *PLK1* y *GADD45A*, en rojo) también se sobreexpresó en las muestras de pacientes pediátricos Hispánicos con LAL-B de la base TARGET.

Por último, la Figura 10 muestra la red de vías, entre ellas Ras y PI3K-Akt, relacionadas a los genes *IL7R* y *PIK3CG* observados en pacientes mexicanos con bajos niveles de expresión.



**Figura 10. Análisis de Vías.** Los GDEs de las muestras de la base TARGET (MO vs SP) se analizaron a través del programa PATHVIEW y se realizó un análisis de enriquecimiento usando la base KEGG. La red de vías resultante muestra las vías Ras y PI3k-Akt relacionadas con los genes *IL7R* y *PIK3CG*, observados en el perfil de expresión global de pacientes mexicanos.

## DISCUSIÓN

Se sabe que en la LAL, las células B progenitoras dejan de diferenciarse, permanecen como linfoblastos inmaduros en la médula ósea y aparecen en los

nódulos linfáticos y el bazo después de la proliferación. Por otro lado, la proliferación celular y la inhibición de la apoptosis son eventos clásicos en la leuquimiogénesis.<sup>22</sup> En concordancia, los resultados de GSEA mostraron un enriquecimiento de genes del ciclo celular (Figura S1), *PLK1*, *CDC20*, *GADD45A*, y *CDC45* mostraron niveles altos de expresión en los linfoblastos pre-B de MO de los pacientes estudiados (Figuras 2 y 5).

*PLK1* tiene funciones importantes en la transición de G2 a fase M y se considera un proto-oncogen porque activa la vía de señalización PI3K/Akt/mTOR en LAL y otras enfermedades oncológicas hematológicas.<sup>23</sup> Su expresión en células leucémicas está regulada a través de una vía dependiente de PI3K- y p38.<sup>24</sup> De acuerdo al avance del ciclo celular mediado por *PLK1*, es posible que la alta expresión de este gen esté favoreciendo la proliferación y la transformación maligna de la célula pre-B. Es de notar que el análisis ARACNE infirió una asociación entre *PLK1* y *OR11H7*, un receptor olfativo expresado en la línea celular HK-2 del túbulo proximal humano.<sup>25</sup>

También se observó un aumento en la expresión de *CDC20*, que activa el complejo/ciclosoma promotor de anafase (APC/C) en el ciclo celular. Un nivel alto de expresión de *CDC20* está significativamente correlacionado con una baja sobrevivencia en la mayoría de los tumores sólidos en humanos.<sup>26</sup> En este estudio se reporta una sobreexpresión de *CDC20* en pre-B ALL, lo cual podría inhibir el arresto mitótico y promover la anafase prematura al desregularizar la activación de APC lo que da como resultado inestabilidad genómica. La aneuploidía altera la proliferación en las células troncales hematopoyéticas; entre los genes involucrados en este proceso están aquellos que regulan el checkpoint mitótico, la

respuesta a daño al DNA, y la recombinación, como podría ser el caso en los pacientes con LAL pre-B de este estudio. De la misma manera, las alteraciones genómicas pueden alterar muchas vías a la vez; por ejemplo, *Simonetti et al.*<sup>27</sup> observaron altos niveles de *CDC20*, *PLK1* y, junto con *RAD50*, una firma de tres genes para la leucemia mieloide aguda (LMA), lo cual indica un proceso de varios pasos que involucra diferentes fases del ciclo celular. Esta evidencia experimental concuerda con los resultados de este análisis para LAL pre-B, ya que los reguladores mitóticos *PLK1* y *CDC20* se sobreexpresaron en linfoblastos pre-B de MO, lo que sugiere que tales alteraciones específicas en el ciclo celular también están presentes en LAL y dan como resultado una proliferación anormal.

El tercer regulador mitótico con altos niveles en el perfil de expresión por microarreglos fue *GADD45A*, asociado *in vivo* con señalización mediada por p38 MAPKs (mitogen-activated protein kinases).<sup>28</sup> Aunque *GADD45A* es uno de varios genes de arresto del crecimiento y pro-apoptóticos, evidencia experimental sugiere que también tiene función en la sobrevivencia de células hematopoyéticas y exhibe un papel anti-apoptótico. *D'Angelo et al.*<sup>29</sup> demostraron que la activación constitutiva *in vivo* de *GADD45A* en blastos leucémicos promueve la sobrevivencia de células hematopoyéticas neoplásicas, lo cual probablemente ocurre via la cinasa p38 y Bcl-xl. La sobreexpresión de *GADD45A* en linfoblastos pre-B de MO en este estudio podría apoyar la sobrevivencia celular a través de vías simultáneas de acuerdo al contexto celular.

El cuarto regulador mitótico fue *CDC45*, que se requiere para la replicación de DNA. *Vaisvilas et al.*<sup>30</sup> reportaron un paciente con una duplicación *de novo* de 6.6 Mb en una región cromosómica que contiene el gen *CDC45*, lo que resultó en



LAL pre-B. Los autores concluyeron que la sobreexpresión de genes, tales como *CDC45*, en la región duplicada responsable del ciclo celular, podrían haber contribuido a la formación del clon leucémico en la médula ósea. Al igual que *PLK1*, *CDC20*, y *GADD45A*, el gen *CDC45* mostró altos niveles de expresión en los linfoblastos pre-B de MO en el análisis de microarreglos. Adicionalmente, el análisis ARACNE infirió una interacción entre *CDC45* y *KCNQ1DN* (SF 5), un lncRNA relacionado con el Tumor de Wilm o nefroblastoma, el cáncer renal pediátrico más común que afecta a niños entre 3-5 años.<sup>31</sup> Aún más, se ha encontrado recientemente que *KCNQ1DN* está notablemente disminuido en el carcinoma de célula renal (RCC, renal cell carcinoma), tejidos y líneas celulares. Reprime el crecimiento celular y progresión del ciclo celular de RCC inhibiendo la expresión de c-Myc. *KCNQ1DN* también inhibe la actividad transcripcional de promotor de c-Myc.<sup>32</sup>

Además de todo esto, el análisis bioinformático comparativo de datos de pacientes pediátricos Hispánicos diagnosticados con LAL-B de la base de datos TARGET, corroboró la sobreexpresión de *PLK1*, *CDC20*, *GADD45A* y *CDC45* y también concuerda con los resultados de *Ma et al.*<sup>33</sup> quienes condujeron un estudio pan-cancer de cánceres pediátricos, incluyendo 689 pacientes con LAL-B. Esto era de esperarse porque, como estos autores encontraron, el ciclo celular es uno de los 21 procesos biológicos con disrupción debida a alteraciones precursoras comunes en los cánceres pediátricos y porque los pacientes pediátricos mexicanos de este estudio fueron indirectamente seleccionados por razones metodológicas y presentaban altos conteos de células blancas (CCBs).

Vistos en conjunto, estos resultados podrían ser útiles para estratificar y subclasificar a los pacientes con LAL-B y, como *Simonetti et al.* sugieren para la LMA, apuntan a la maquinaria mitótica como un blanco terapéutico potencial.

El análisis de expresión global por microarreglos también reveló que entre los genes con bajos niveles de expresión en los linfoblastos pre-B de MO, *IL7R* y *PIK3GC* (Figuras 2 y 5) podrían explicar la inhibición de la diferenciación de la célula B.

La señalización de *IL7R* activa la señalización de JAK/STAT y las vías de PI3K/Akt/mTOR, que están entre las alteraciones precursoras histotipo-específicas en leucemias<sup>33</sup>, y es necesaria para el desarrollo normal y mantenimiento de células B y T y crucial para la leuquimogénesis; se han descrito mutaciones activadoras de ganancia-de-función en LAL-B. Sin embargo, en humanos, las mutaciones inactivadoras de *IL7R* dan como resultado linfopenia severa de células T, con números normales, aunque no funcionales de células B.<sup>34</sup> Además de esto, la gráfica del RNAm de expresión (RNAseq) de *IL7R* en la Cancer Cell Line Encyclopedia (CCLE, por sus siglas en inglés) muestra un valor de expresión bajo.<sup>35</sup> Esto concuerda con este estudio ya que *IL7R* exhibió casi el mismo nivel de expresión bajo reportado por la CCLE. Aunque existen vías alternativas para activar PI3K/Akt/mTOR, una baja en la expresión de *IL7R* podría desregular el desarrollo de célula B y la función apropiada de estas vías.

El segundo gen más significativamente subexpresado fue *PIK3CG*. Es necesario mencionar primero que la señalización de PI3K regula numerosos procesos biológicos, así que su baja expresión da como resultado una proliferación celular baja y un aumento de la muerte celular.<sup>36</sup> De igual manera, la

inhibición farmacéutica de la vía PI3K/AKT lleva a la baja de la proliferación celular en la LAL pre-B.<sup>37</sup> La vía de señalización PI3K puede ser activada por receptores acoplados a proteína G (GPCR, por sus siglas en inglés) a través de la interacción directa con dímeros de *PIK3CG* (isoforma catalítica PI3K p110 gamma) y proteínas Ras, que están implicadas en varios aspectos de la función y regulación inmune. Adicionalmente, *PIK3CG* se requiere para el desarrollo temprano de la célula B y contribuye a su mantenimiento, proliferación y transformación.<sup>38</sup> Así, la baja expresión de *PIK3CG* e *IL7R* podría explicar en parte el arresto del desarrollo normal de las células B en los linfoblastos pre-B de MO.

Por otra parte, es de notar que la resistencia a la quimioterapia anticancer ocurre en parte porque el ambiente de la médula ósea promueve la adhesión celular mediada por resistencia a medicamentos en células leucémicas y dispara la señalización pro-sobrevivencia (p.ej., la vía PI3K/AKT) que permite que las células leucémicas soporten la quimioterapia. Si *PIK3CG* está efectivamente involucrado en la resistencia a quimioterapia en células leucémicas, es razonable hipotetizar que los bajos niveles de expresión observados en este análisis, no generarían resistencia a quimioterapia. De todos los pacientes incluidos en este estudio, sólo dos fallecieron, mientras que el resto no parecen haber tenido linfoblastos pre-B resistentes y sobrevivieron.

Interesantemente, el análisis bioinformático comparativo de pacientes Hispánicos con LAL-B de la base de datos TARGET corroboraron la subexpresión de *IL7R*, la cual está considerada una mutación precursora en común del cáncer<sup>33,39</sup>, pero no la de *PIK3CG*. Los resultados mostraron que estos pacientes de LAL-B expresaban bajos niveles de *PIK3CA*, *PIK3CB* y *PIK3CD*, que son PI3Ks

Clase I, tales como *PIK3CG*, *PIK3R1* y *PIK3R2*. En el mismo sentido, *Ma et al.* observaron a *PIK3CA* y *PIK3R1* como las mutaciones más frecuentes en las vías PI3K y RAS en leucemias. En contraste, *Gyurina et al.* reportaron la expresión relativa de *PIK3CG* en linfoblastos pre-B de MO de LAL y señalaron que la vía PIK3/AKT juega un papel regulatorio principal en LAL-BCP.<sup>33,40</sup>

El gen *AIDA* también mostró niveles bajos en el perfil de expresión de microarreglos. Esta proteína se localiza en el citosol, los microtúbulos y el puente citocinético. Los dominios C2 en las proteínas AIDA son importantes para entender la interacción entre el esqueleto microtubular y de microfilamentos y las membranas celulares. La calmodulina es un dominio de interacción citoesquelético asociado a los dominios C2 de *AIDA*. Estos hallazgos relacionan a *AIDA* con *CALN1*, otro gen validado en este estudio (Tabla 2) que posee dominios C2 como *AIDA*, implicando la posible interacción de ambos genes en las funciones del citoesqueleto y el tráfico de vesículas en LAL pre-B.

Sorpresivamente, el análisis ARACNE reveló dos genes asociados con *AIDA* (*STAP-2* y *FGF19*) los cuales están involucrados en la vía JAK/STAT que regula positivamente la división celular, sobrevivencia, diferenciación y el sistema inmune.<sup>41</sup>

El primer gen, *STAP-2*, y el “hub” de la red inferida por ARACNE, codifica para una proteína adaptadora que modula la actividad transcripcional de STAT3 y STAT5.<sup>42</sup> *STAP-2* se expresa en linfocitos y juega un papel crucial en el sistema inmune al controlar la transducción de señal de citosinas y modular tanto al sistema inmune innato como al adaptativo.<sup>43</sup>

El segundo gen, *FGF19*, tiene una alta afinidad por el Receptor al Factor de crecimiento de fibroblastos 4 (FGFR4). Los FGFRs se requieren para las actividades biológicas de FGF; una vez activados, fosforilan residuos específicos de tirosina que median la interacción con las vías JAK/STAT, PI3K-AKT y RAS-MAPK.<sup>44</sup> *Su et al.*<sup>45</sup> observaron que los niveles de *FGF-19* en pacientes de LMA antes de la quimioterapia eran significativamente más altos que los del grupo control. Aunque *FGF19* no estuvo entre los GDEs del análisis de expresión global, y por tanto, no fue validado por RT-qPCR, los resultados de ARACNE apuntan a que podría estar involucrado en la LAL pre-B.

*SMIM10L2B*, el tercer gen en la red ARACNE, es un RNA largo no-codificante (lncRNA). Los lncRNAs participan en la diferenciación normal de la célula B, pero su desregulación está involucrada en el desarrollo de malignidades de la célula B, como la LAL. El perfil de expresión de los lncRNAs durante el desarrollo de la célula B se ha realizado en varios estudios que reportan patrones de expresión específicos para el tipo celular en varios estados del desarrollo de la célula B.<sup>46-47</sup> La observación del lncRNA *SMIM10L2B* en este análisis es sorprendente porque aún no ha sido reportado en el tejido de MO o como parte de la firma de lncRNAs involucrados en LAL-B observados en estudios recientes.<sup>48</sup> Se requerirá más investigación para entender mejor su papel y significancia en este tipo de cáncer sanguíneo.

## CONCLUSIONES

1. El perfil de expresión génica global de linfoblastos pre-B CD10+/CD19+ de MO y SP purificados por FACS y el análisis bioinformático mediante TAC, GSEA e IPA revelaron un firma de cuatro genes reguladores mitóticos (*PLK1*, *CDC20*, *GADD45A* y *CDC45*) que promueve la proliferación de estas células leucémicas y una firma de dos genes con baja expresión (*IL7R* and *PIK3CG*) que explica en parte la inhibición de la diferenciación de los linfoblastos pre-B en MO.
2. Aunque el tamaño de muestra de este estudio fue pequeño, ya que los pacientes con bajos CCBs no fueron elegibles para el protocolo FACS que se implementó, sí proveyó muestras puras de linfoblastos pre-B CD10+/CD19+ de MO y SP sin otras células blancas o no leucémicas contaminantes.
3. Es de notar que *ETV6-RUNX1*, una alteración promotora de LAL<sup>49</sup> fue el único rearrreglo cromosómico detectado, y que el paciente que lo portaba eventualmente falleció. *Bekker et al.*<sup>6</sup> determinaron su prevalencia (7.4%) en la médula ósea de 240 pacientes pediátricos mexicanos con LAL-B y sugieren que la agresividad de la LAL en niños mexicanos pudiera deberse a la ocurrencia de este y otros rearrreglos génicos importantes. El número de muestras es otra posible explicación a la baja presencia de *ETV6-RUNX1* y los otros tres rearrreglos cromosómicos.
4. La firma de cuatro genes del ciclo celular observada en el perfil de expresión global fue corroborada a través de un análisis comparativo de muestras de la base de datos TARGET. Aunque este importante proceso tiene muchos otros genes (>200), los cuales también se sobreexpresan en LAL, esta firma génica parece ser importante, ya que tres de los genes han sido observados también en LMA; aquí

se reporta su sobreexpresión en diferentes puntos del ciclo celular en el contexto de LAL pre-B.

5. En conjunto, esta evidencia experimental podría usarse para estratificar y subclasificar pacientes con LAL-B con altos CCBs y, como se ha sugerido en el caso de la LMA con *CDC20* y *PLK1*, señala a estas proteínas de la maquinaria mitótica (*PLK1*, *CDC20*, *CDC45* y *GADD45A*) como blancos terapéuticos potenciales en la LAL pre-B.

6. De los dos genes subexpresados, *IL7R* fue corroborado por el análisis bioinformático comparativo con muestras de pacientes con LAL-B de la base TARGET, pero *PIK3CG* no. Se propone que los bajos valores de expresión de *PIK3CG* podrían ser un rasgo característico de los pacientes pediátricos Mexicanos con LAL-B y subraya la importancia de la vía PI3K-mTOR como blanco terapéutico.

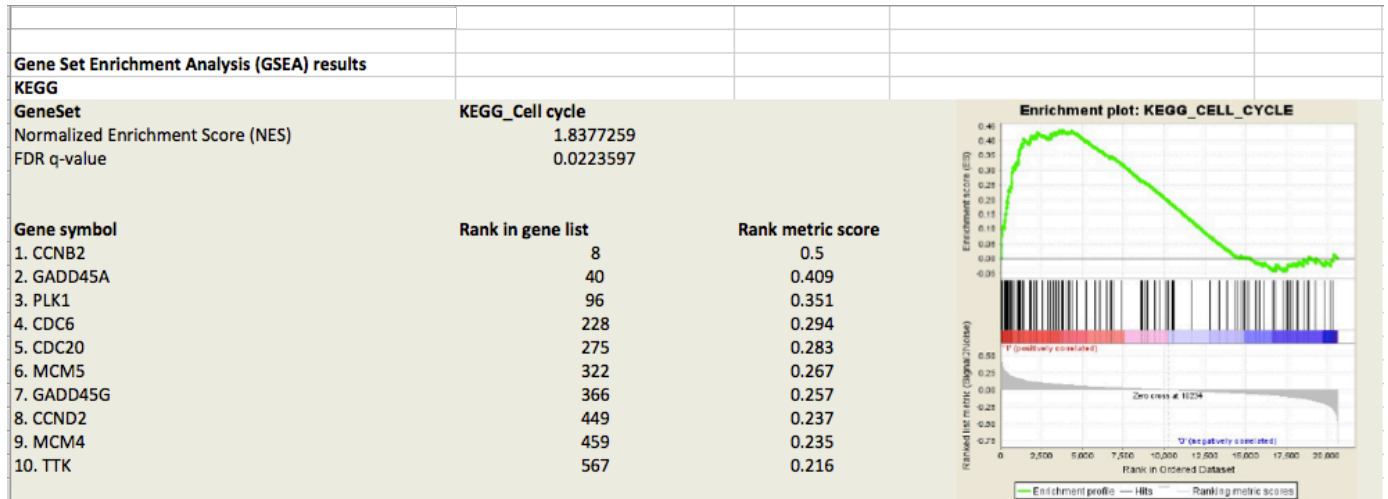
7. Se encontraron similitudes y diferencias entre el grupo de pacientes Mexicanos estudiado y los pacientes Hispánicos de la base TARGET.

Las razones para la alta incidencia, recaída temprana y baja tasa de sobrevivencia de los pacientes pediátricos con LAL-B merece más investigación y podría explicarse no sólo por características genéticas sino también por factores epigenéticos controlados por variables ambientales.

\* Los datos discutidos en este estudio han sido depositados en el repositorio de NCBI, Gene Expression Omnibus (Edgar *et al.*, 2002)<sup>50</sup> y están disponibles a través del *GEO Series accession number* GSE168593 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE168593>).

\* The data discussed in this study have been deposited in NCBI's Gene Expression Omnibus (Edgar *et al.*, 2002)<sup>50</sup> and are accessible through GEO Series accession number GSE168593 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE168593>

## APÉNDICE de FIGURAS SUPLEMENTARIAS



**Figura S1.** Resultados de GSEA (Gene Set Enrichment Analysis).



## Ingenuity Pathway Analysis (IPA) results

## Top Canonical Pathways

Name	p-value
Cell cycle: G2/M DNA Damage Checkpoint Regulation	1.98E-04
Cell cycle control of chromosomal replication	3.98E-03

## Diseases and Disorders

Name	p-value	Molecules
Cancer	4.60E-02 - 1.07E-06	94
Hematological Disease	4.60E-02 - 1.07E-06	42
Immunological Disease	4.60E-02 - 1.07E-06	49

## Molecular and cellular functions

Name	p-value	Molecules
Cellular movement	4.38E-02 - 3.85E-05	21
Cellular growth and Proliferation	3.86E-02 - 3.45E-04	12
Cell Cycle	4.62E-02 - 1.45E-03	13

## Physiological system development and function

Name	p-value	Molecules
Hematological system development and function	4.38E-02 - 3.85E-05	11
Immune cell trafficking	4.38E-02 - 3.85E-05	9

## Top Networks

Networks	Networks
1. Hematological System Development and Function, Lymphoid Tissue Structure and Development, Tissue, Morphology	47
2. Hematological Disease, Immunological Disease, Cell Cycle	28
3. Cellular Movement, Cell-To-Cell Signaling and Interaction, Organismal Development	19
4. Cellular Movement, Immune Cell Trafficking, Hematological System Development and Function	19
5. Developmental Disorder, Hereditary Disorder, Neurological Disease	2

## Top Analysis-Ready Molecules

## Expr Fold Change Up-regulated

Molecules	Expr. Value
CNN3	4.809
EDIL3	4.652
LHFPL2	4.547
UBE2E3	3.775
STON2	3.724
CALN1	3.427
MTSS1L	2.645
SPNS3	2.576
ZNF462	2.561
MXD1	2.463

## Expr Fold Change Down-regulated

Molecules	Expr. Value
SKAP2	-5.724
CAMK2D	-5.301
NT3E	-5.128
IL7R	-3.95
RPL31P57	-3.318
PIK3CG	-3.161
GPR63	-2.957
CD86	-2.926
PHF11	-2.841
EZR	-2.577

Figura S2. Resultados de IPA (Ingenuity Pathway Analysis).

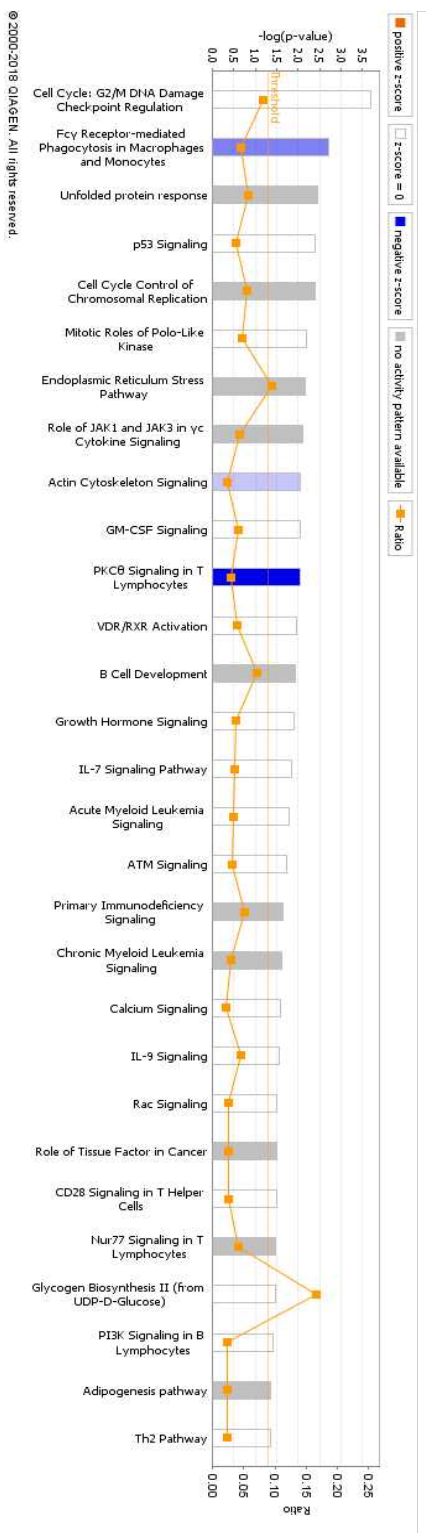
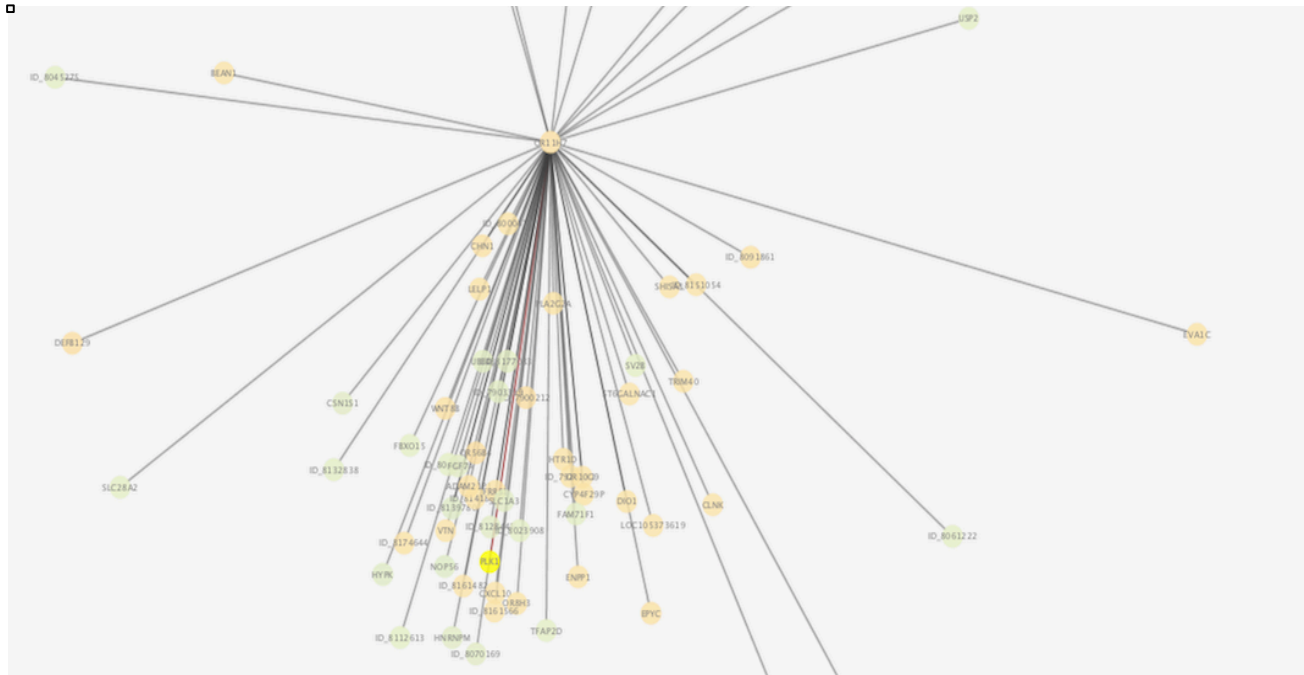
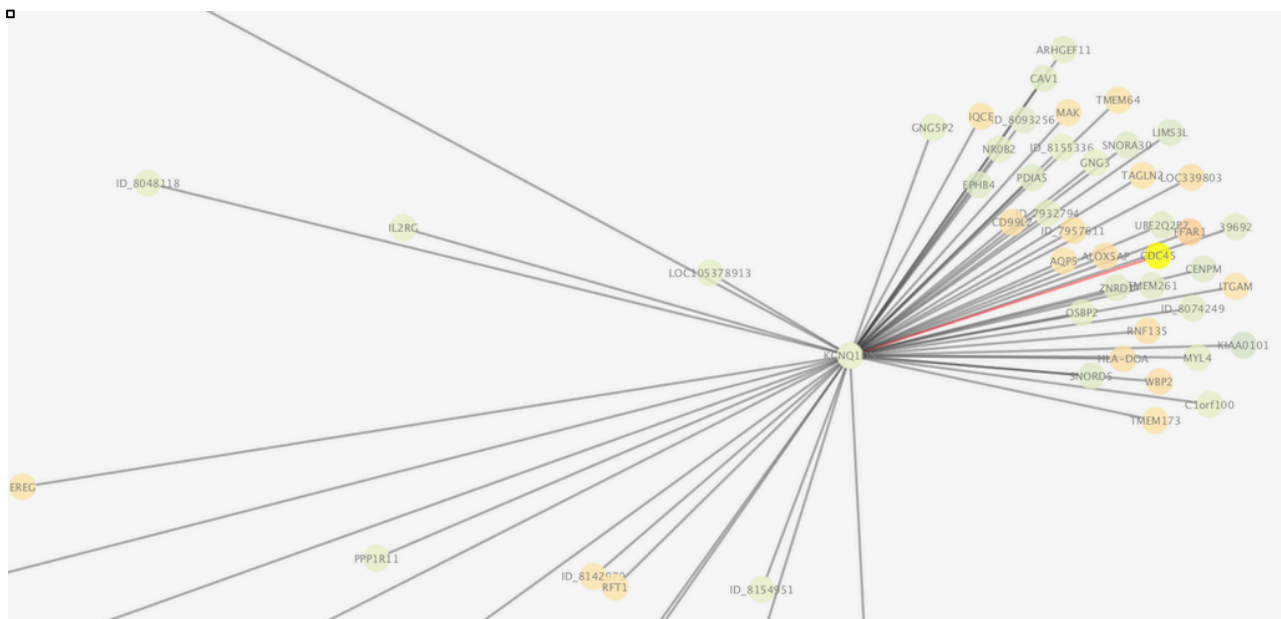


Figura S3. Diagrama de resultados de IPA (Ingenuity Pathway Analysis).



**Figura S4.** Red ARACNE (MI 07) mostrando la asociación entre el gen del ciclo celular *PLK1* y el gen de receptor olfatorio *OR11H7*.



**Figura S5.** Red ARACNE (MI 07) mostrando la asociación entre el gen del ciclo celular *CDC45* y *KCNQ1DN*, un gen de la clase lncRNA.

## ÁPÉNDICE DE PROTOCOLOS DE TRABAJO

**PROTOCOLO 1.** Separación de células mononucleares de sangre periférica (SP) y médula ósea (MO) por Histopaque-1077.

**PROTOCOLO 2.** Tinción de células mononucleares para inmunotipificación de linfoblastos pre-B CD10+/CD19+ mediante FACS.

**PROTOCOLO 3.** Extracción de RNA total por Trizol.

**PROTOCOLO 4.** Microarreglos (Unidad de Microarreglos, INMEGEN).

**PROTOCOLO 5.** RT-PCR del RNA total de muestras de SP y MO de pacientes pediátricos diagnosticados con LAL-B.

**PROTOCOLO 6.** RT-qPCR del cDNA de RNA total de muestras de SP y MO de pacientes pediátricos diagnosticados con LAL-B, para validación de los 26 principales GDEs significativos: *PLK1*, *CDC20*, *CALN1*, *PTPN14*, *HIST1H2BI*, *HIST1H3F*, *HIST1H3G*, *CDC45*, *STON2*, *CNN3*, *RUNX2*, *IL7R*, *TCL1A*, *CD72*, *HDAC*, *NT5E*, *CD86*, *GPR65*, *LILRB2*, *LY86*, *AIDA*, *CCNB2*, *DOCK1*, *GADD45A*, *PIK3CG*, *HCK*.

**PROTOCOLO 1. Separación de células mononucleares de sangre periférica (SP) y médula ósea (MO) por Histopaque-1077.**

Después de asistir al Servicio Oncológico correspondiente, el mismo día de la colecta de la muestra de entre 3-5 ml de SP o MO del paciente en tubos Vacutainer con EDTA como anticoagulante, y antes de que pasaran 48 h, se separaron las células mononucleares por Histopaque-1077 mediante el siguiente protocolo.

1. En un tubo de centrifuga cónico de 15 ml añadir 3 ml de Histopaque-1077 y dejar que llegue a temperatura ambiente.
2. Depositar cuidadosamente sobre el Histopaque-1077, 3 ml de la muestra (SP o MO).
3. Centrifugar a 400'g/ 30 minutos exactamente y a temperatura ambiente.
4. Aspirar cuidadosamente la fase superior con una pipeta Pasteur hasta 0.5 cms de la interfase opaca que contiene las células mononucleares. Descartar la fase superior.
5. Transferir cuidadosamente la interfase opaca con una pipeta Pasteur a un tubo de centrifuga cónico limpio.
6. Lavar las células añadiendo 10 ml de PBS y mezclar suavemente introduciendo y retirando la pipeta Pasteur.
7. Centrifugar a 250'g/ 10 minutos.
8. Aspirar el sobrenadante y descartar.
9. Resuspender el pellet de células con 5 ml de PBS y mezclar suavemente introduciendo y retirando la pipeta Pasteur.
10. Centrifugar a 250'g/ 10 minutos.
11. Repetir los pasos 8, 9 y 10, descartar el sobrenadante y resuspender el pellet de células en 1 ml de PBS.

## **PROTOCOLO 2. Tinción de células mononucleares para la inmunotipificación de linfoblastos pre-B CD10+/CD19+ mediante FACS.**

Inmediatamente después de la separación de células mononucleares de cada muestra de sangre periférica (SP) o médula ósea (MO) por Histopaque-1077, se prepararon tubos para la tinción de estas para la inmunotipificación de linfoblastos pre-B CD10+/CD19+ mediante FACS en el siguiente orden:

1. Tubo “blanco” con muestra sin anticuerpos: para compensación y ubicación de la población celular en la adquisición de datos.
2. Tubo con muestra teñida con anticuerpo CD10 acoplado al fluorocromo FITC (Fluoroceín-Iso-Tio-Cianato) para compensación.
3. Tubo con muestra teñida con anticuerpo CD19 acoplado al fluorocromo APC (Alofocianina) para compensación.
4. Tubo con muestra teñida con anticuerpo CD45 acoplado al fluorocromo PB (Pacific Blue) para compensación.
5. Tubo con muestra teñida con anticuerpo CD34 acoplado al fluorocromo PE (Ficoeritrina) para compensación.
6. Tubo de muestra total teñida con anticuerpos: CD10-FITC, CD19-APC, CD45-PB y CD34-PE para el pre-sort para “llamar” a todos los colores de los fluorocromos, y así, a las células que se unieron.

<i><b>Muestra SIN Anticuerpos acoplados a fluorocromos</b></i>	<i><b>Muestra + Anti CD10 acoplado a FITC</b></i>	<i><b>Muestra + Anti CD19 acoplado a APC</b></i>	<i><b>Muestra + Anti CD45 acoplado a PB</b></i>	<i><b>Muestra + Anti CD34 acoplado a PE</b></i>	<i><b>Muestra Total + Anti CD10- FITC Anti CD19-APC Anti CD45-PB Anti CD34-PE</b></i>
<b>1 mL PBS</b>	<b>1 mL PBS</b>	<b>1 mL PBS</b>	<b>1 mL PBS</b>	<b>1 mL PBS</b>	<b>1 mL PBS</b>
<b>1 uL Muestra</b>	<b>1 uL Muestra</b>	<b>1 uL Muestra</b>	<b>1 uL Muestra</b>	<b>1 uL Muestra</b>	<b>Muestra Total</b>
	<b>1 uL Anti CD10-FITC</b>	<b>1 uL Anti CD19-APC</b>	<b>1 uL Anti CD45-PB</b>	<b>1 uL Anti CD34-PE</b>	<b>1 uL CD10-FITC 1uL CD19-APC 1 uL CD45-PB 1 uL CD34-PE</b>

Los tubos se incubaron por 30 minutos a 4°C y luego se analizaron y purificaron por citometría de flujo activada por fluorescencia (FACS, fluorescence activated cell sorting) en un citómetro de flujo FACS Aria II (BD Biosciences, San Jose, CA, USA). Los linfocitos fueron identificados por sus propiedades *side scatter* y *forward scatter* y se seleccionaron los CD45<sup>+</sup> CD10<sup>+</sup> CD34<sup>+</sup> CD19<sup>+</sup> para obtener finalmente los linfoblastos pre-B CD10<sup>+</sup>/CD19<sup>+</sup>. La pureza fue de 90-95% para todas las muestras que quedaron suspendidas en PBS.

### **PROTOCOLO 3. Extracción de RNA total por TRIzol.**

Inmediatamente después de la inmunotipificación de los linfoblastos pre-B CD10+/CD19+ con una pureza de 90-100% mediante FACS, las muestras se centrifugaron a 5,000 rpm/ 3 min y se lavaron con PBS antes de la extracción de RNA total por el método de TRIzol.

#### **Lisado de muestras y separación de fases.**

1. Lisar y homogenizar la muestra en TRIzol de acuerdo con el material de inicio, en este caso:

- Células en suspensión.

a. Colectar las células por centrifugación y descartar el sobrenadante.

b. Añadir 0.75 mL of TRIzol por cada 0.25 mL de muestra ( $5-10 \times 10^6$  células de origen animal) a el pellet.

Nota: No lavar las células antes de añadir TRIzol para evitar degradación de RNAm.

c. Pipetear el lisado arriba y abajo varias veces para homogenizar.

Nota: El volumen de la muestra no deberá exceder 10% del volumen de TRIzol usado para la lisis.

\* Hasta aquí, las muestras se pueden almacenar a 4°C toda la noche o a 20°C hasta por 1 año. Si no:

2. Incubar por 5 minutos para permitir la disociación completa del complejo de nucleoproteínas.

3. Añadir 0.2 mL de cloroformo por 1 mL de TRIzol usado para la lisis, cerrar el tubo y mezclar completamente por agitación.

4. Incubar por 2-3 minutos.

5. Centrifugar la muestra por 15 min/  $12,000 \times g$  a 4°C.

La mezcla se separa en una interfase inferior fenol-cloroformo roja, y una fase superior incolora acuosa.

6. Transferir la fase acuosa que contiene el RNA a un tubo nuevo angulando el tubo a 45° y pipeteando la solución hacia fuera.



**Aislamiento de RNA**

- a. Añadir 0.5 mL de isopropanol a la fase acuosa por 1 mL de TRIzol usado para lisis.
- b. Incubar por 10 minutos a 4°C.
- c. Centrifugar por 10 min/ 12,000 × g a 4°C. El precipitado de RNA total forma un pellet tipo gel blanco en el fondo del tubo.
- d. Descartar el sobrenadante con una micropipeta.

**1. Precipitación del RNA**

- a. Resuspender el pellet en 1 mL de alcohol 75% por 1 mL de TRIzol usado para lisis.

Nota: El RNA puede ser almacenado en alcohol 75% por al menos 1 año a -20°C, o al menos 1 semana a 4°C.

- b. Agitar la muestra por vortex brevemente, luego centrifugar 5 min/ 7,500 × g a 4°C.
- c. Descartar el sobrenadante con una micropipeta.
- d. Aspirar o secar al aire la pellet de RNA por 5–10 minutos.

IMPORTANTE: No secar el pellet con centrifuga de vacío. No dejar secar el pellet de RNA, para asegurar la solubilización total del RNA. Cuando las muestras de RNA están parcialmente disueltas, tienen un radio  $A_{230}/280 < 1.6$ .

**2. Lavado del RNA**

- a. Resuspender el pellet en 20–50 µL de agua libre de RNasa pipeteando hacia arriba y hacia abajo.
- b. Incubar en baño María a 55–60°C por 10–15 minutos.

Las muestras con un radio  $A_{260}/280 \geq 1.8$  fueron seleccionadas y almacenadas a -70°C hasta su uso. La integridad del RNA (RIN: RNA Integrity Number) se evaluó con el Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). Sólo las muestras con un RIN >7.5 se incluyeron en el estudio.

**PROTOCOLO 4.** Microarreglos (Unidad de Microarreglos, INMEGEN).

La integridad del RNA (RIN) de cada muestra se evaluó por electroforesis de capilaridad usando un Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA). Sólo las muestras con un RIN >6.0 se incluyeron en el análisis por microarreglos (AM).

El MA se realizó de acuerdo a los protocolos del fabricante (Affymetrix Inc., Santa Clara, CA). Brevemente, se sintetizó el cDNA de primera cadena a partir de 150 ng de RNA total y usando la transcriptasa reversa Superscript II unida a un oligomero poly (T) oligomer que incorporó el promotor T7. Para obtener cRNA, la síntesis de cDNA de segunda cadena fue seguida por transcripción *in vitro*. Este cRNA se uso como plantilla para un segundo ciclo de síntesis de cDNA con dUTPs que se incorporaron en la nueva cadena. Después de la fragmentación del cDNA usando uracil-DNA glicosilasa y purin-pirimidin endonucleasa, los fragmentos fueron marcados mediante reacción de adición de deoxinucleótido terminal marcado con biotina.

El producto de cDNA marcado se calentó a 95°C y se hibridó al microarreglo Human Gene 1.0 ST (Affymetrix Inc.) por 17 horas a 45°C. Las muestras se lavaron con buffers de baja y alta astringencia y se tiñeron con estreptavidin-ficoeritrina usando la Affymetrix Fluidics Station 450. El GeneChip Scanner 3000 7G (Affymetrix, Inc) se usó para colectar las señales fluorescentes y el software Expression Console (Affymetrix, Inc) se utilizó para obtener la señal de intensidad (intensity signal) y datos de calidad (quality data) de los microarreglos escaneados. El software proporciona informes resumidos y mediciones incluyendo fondo promedio, área bajo la curva positiva vs negativa y radios de expresión para los picos en los controles, útiles para la evaluación QC del microarreglo. El método global de escalamiento, o normalización se aplico a todos los GeneChips. Las medidas de calidad e igualmente el porcentaje de genes presentes y el radio de genes endógenos, indicaron una alta calidad total de muestras y ensayos.

**PROTOCOLO 5.** RT-PCR del RNA total de muestras de SP y MO de pacientes pediátricos diagnosticados con LAL-B.

Primero se hicieron pruebas para estandarizar las condiciones para sintetizar el cDNA a partir del RNA total de cada muestra (SP = 9 y MO = 4).

Se utilizó el kit **First Strand cDNA Synthesis** siguiendo el protocolo del fabricante (Thermo Scientific, USA).

**I. Síntesis de cDNA de la Primera Hebra.**

Después de descongelar, mezclar los componentes del kit. Almacenar/ poner en hielo.

1. Añadir los siguientes reagentes en un tubo estéril, libre de nucleasas, sobre hielo en el orden indicado:

<b>MIX 1</b>	<b>1X</b>	<b>3X (Reacciones)</b>	<b>13X (Muestras)</b>
<b>Random hexamer primer</b>	1 $\mu$ l	3 $\mu$ l	39 $\mu$ l
<b>Agua ultrapura, libre de nucleasas</b>	8 $\mu$ l	24 $\mu$ l	
		27 $\mu$ l	351 $\mu$ l
			390 $\mu$ l/13 muestras = 30 $\mu$ l/tubo
<b>RNA total</b>	1 $\mu$ l	3 $\mu$ l	<b>33<math>\mu</math>l/tubo</b>

2. Colocar los tubos a 65-70°C por 10 minutos y después pasar a hielo.

3. Añadir los siguientes componentes en el orden indicado:

<b>MIX 2</b>	<b>1X</b>	<b>3X (Reacciones)</b>	<b>13X (Muestras)</b>
<b>5X</b> (Buffer de reacción)	4 $\mu$ l	12 $\mu$ l	156 $\mu$ l
<b>Inhibidor</b> (RiboLock RNase inhibitor)	1 $\mu$ l	3 $\mu$ l	39 $\mu$ l
<b>dNTP Mix</b> (10 mM)	2 $\mu$ l	6 $\mu$ l	78 $\mu$ l
<b>Transcriptasa reversa</b> (M-MuLV Reverse Transcriptase)	2 $\mu$ l	6 $\mu$ l	78 $\mu$ l
			351 $\mu$ l/13 muestras = 27 $\mu$ l/tubo

4. Colocar los tubos a 42°C por 2 horas.
5. Colocar los tubos a 65-70°C por 5 minutos y después pasar a hielo.
6. Almacenar a -20°C.

## **II. Amplificación por PCR de cDNA de la primera hebra.**

El producto de la síntesis de cDNA de la primera hebra puede ser usado directamente en PCR o qPCR.

**PROTOCOLO 6.** RT-qPCR del cDNA de RNA total de muestras de SP y MO de pacientes pediátricos diagnosticados con LAL-B, para validación de los 26 principales GDEs significativos:

*PLK1, CDC20, CALN1, PTPN14, HIST1H2BI, HIST1H3F, HIST1H3G, CDC45, STON2, CNN3, RUNX2, IL7R, TCL1A, CD72, HDAC, NT5E, CD86, GPR65, LILRB2, LY86, AIDA, CCNB2, DOCK1, GADD45A, PIK3CG, HCK.*

Para los experimentos de RT-qPCR se utilizó el kit Maxima SYBER Green/ROX qPCR Master Mix (2X) siguiendo las instrucciones del fabricante (Thermo Scientific, USA).

### I. Preparación de la reacción.

1. Agitar suavemente en vortex y centrifugar brevemente todas las soluciones después de descongelar.
2. Preparar una mezcla maestra de reacción (*reaction master mix*) añadiendo los siguientes componentes (excepto el cDNA) a un tubo a temperatura ambiente.
3. Mezcle la mezcla maestra totalmente y coloque los volúmenes apropiados en tubos de PCR.
4. Añadir el cDNA a cada tubo de PCR individual que contienen la mezcla maestra.

1X		$\beta$ 2m 13X Muestras	GEN 1 13X Muestras	GEN 2 13X Muestras	GEN 26 13X Muestras
<b>Mezcla maestra (2X)</b> (Maxima SYBER Green/ROX qPCR)	<b>10<math>\mu</math>l</b>	130 $\mu$ l	130 $\mu$ l	130 $\mu$ l	130 $\mu$ l
<b>P1 (Fwd)</b> <b>P2(Rev)</b> (Oligos validados)	<b>1<math>\mu</math>l</b> <b>1<math>\mu</math>l</b>	13 $\mu$ l 13 $\mu$ l	13 $\mu$ l 13 $\mu$ l	13 $\mu$ l 13 $\mu$ l	13 $\mu$ l 13 $\mu$ l
<b>Agua ultrapura,</b> libre de nucleasas	<b>6<math>\mu</math>l</b>	78 $\mu$ l	78 $\mu$ l	78 $\mu$ l	78 $\mu$ l
		234 $\mu$ l/13 = <b>18<math>\mu</math>l</b>	234 $\mu$ l/13 = <b>18<math>\mu</math>l</b>	234 $\mu$ l/13 = <b>18<math>\mu</math>l</b>	234 $\mu$ l/13 = <b>18<math>\mu</math>l</b>
<b>cDNA</b>	<b>1.8<math>\mu</math>l</b>	<b>19.8<math>\mu</math>l</b>	<b>19.8<math>\mu</math>l</b>	<b>19.8<math>\mu</math>l</b>	<b>19.8<math>\mu</math>l</b>

\*  $\beta$ 2m: beta-2 mioglobulina, control interno.

Para los qPCRs de los 26 GDEs seleccionados se repitieron los pasos 2 a 4 para 4 mezclas maestras que se usaron en 3 experimentos independientes para 7 genes cada uno (más el control interno  $\beta$ 2m en cada caso) y 1 mezcla maestra para un experimento independiente para 5 genes (más el control interno  $\beta$ 2m).

5. Mezclar las reacciones suavemente sin crear burbujas (no usar vortex). Centrifugar brevemente si es necesario.

6. Programar el termociclador de acuerdo a los valores estandarizados, colocar las muestras en el ciclador y comenzar el programa.

<b>STAGE 1</b>	95°C/ 10 min
<b>STAGE 2</b>	<b>50 ciclos</b> 94°C/ 40 min 58°C/ 1 min 72°C/ 40 seg
<b>STAGE 3</b> <b>Curva de disociación</b>	1 repetición 95°C/ 15 seg 60°C/ 30 seg 95°C/ 15 seg

## REFERENCIAS

1. Inaba H, Mullighan CG. Pediatric acute lymphoblastic leukemia. *Haematologica*. 2020;105:2524-39. <https://doi.org/10.3324/haematol.2020.247031>.
2. Jiménez-Hernández E, Jaimes-Reyes EZ, Arellano-Galindo J, García-Jiménez X, Tiznado-García HM, Dueñas-González MT, 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. <https://doi.org/10.1155/2015/940278>.
3. GLOBOCAN. March 2021. <https://gco.iarc.fr/today/data/factsheets/populations/484-mexico-fact-sheets.pdf>
4. Pérez-Saldivar ML, Fajardo-Gutiérrez A, Bernáldez-Ríos R, Martínez-Avalos A, Medina-Sanson A, Espinosa-Hernández L, et al. Childhood acute leukemias are frequent in Mexico City: descriptive epidemiology. *BMC Cancer*. 2011;11:355-65. <https://doi.org/10.1186/1471-2407-11-355>.
5. Moreno-Estrada A, Gignoux CR, Fernandez-Lopez JC, Zakharia F, Sikora M, Contreras AV, et al. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science*. 2014;344:1280–1285. <https://doi.org/10.1126/science.1251688>.

6. Staal FJT, van der Burg M, Wessels, LFA, Barendregt BH, Baert, MRM, van den Burg CMM et al. DNA microarrays for comparison of gene expression profiles between diagnosis and relapse in precursor-B acute lymphoblastic leukemia: choice of technique and purification influence the identification of potential diagnostic markers. *Leukemia*. 2003; 17:1324–1332.

<https://doi.org/10.1038/sj.leu.2402974>

7. Bekker-Méndez VC, Miranda-Peralta E, Núñez-Enríquez JC, Olarte-Carrillo I, Guerra-Castillo FX, Pompa-Mera EN, 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;2014:1–8.

<https://doi.org/10.1155/2014/210560>.

8. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res*. 2016; 44:14753. <https://doi.org/10.1093/nar/gkw419>.

9. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera RD, et al. ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*. 2006;7(S1):S7.

<https://doi.org/10.1186/1471-2105-7-S1-S7>.

10. Guitart-Pla O, Kustagi M, Rugheimer F, Califano A, Schwikowski B. The Cyni



framework for network inference in Cytoscape. *Bioinformatics*. 2015;31:1499–501.

<https://doi.org/10.1093/bioinformatics/btu812>.

11. cBio. Cancer Genomics Data Portal. <https://cbio.mskcc.org/tools/cancer-genomics/index.html>

12. cBioPortal. <https://www.cbioportal.org>

13. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;2:6(269):pl. <https://doi.org/10.1126/scisignal.2004088>.

14. National Cancer Institute Office of Cancer Genomics. Data Access. <https://ocg.cancer.gov/data/data-access>

15. National Cancer Institute Office of Cancer Genomics. Target: Overview. 2020. <https://ocg.cancer.gov/programs/target/overview>

16. Ge SX, Son EW, Yao R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics*. 2018;19:534. <https://doi.org/10.1186/s12859-018-2486-6>.

17. Jiménez-Jacinto V, Sanchez-Flores A, Vega-Alvarado L. Integrative Differential Expression Analysis for Multiple EXperiments (IDEAMEX): a web server tool for

integrated RNA-seq data analysis. *Front Genet.* 2019;10:279.  
<https://doi.org/10.3389/fgene.2019.00279>.

18. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics.* 2020;36:2628-9.  
<https://doi.org/10.1093/bioinformatics/btz931>.

19. Luo W, Pant G, Bhavnasi YK, Blanchard Jr SG, Brouwer C. Pathview Web: userfriendly pathway visualization and data integration. *Nucleic Acids Res.* 2017;45(W1):W501-8. <https://doi.org/10.1093/nar/gkx372>.

20. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert EL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102:15545-50.  
<https://doi.org/10.1073/pnas.0506580102>.

21. Knaack SA, Siahpirani AF, Roy S. A pan-cancer modular regulatory network analysis to identify common and cancer-specific network components. *Cancer Inform.* 2014;13:69-84. <https://doi.org/10.4137/CIN.S14058>.

22. Reyes-Sebastian J, Montiel-Cervantes LA, Reyes-Maldonado E, Dominguez-Lopez ML, Ortiz-Butron R, Castillo-Alvarez A, et al. Cell proliferation and inhibition of apoptosis are related to c-Kit activation in leukaemic lymphoblasts. *Hematology.* 2018; 23:486–95. <https://doi.org/10.1080/10245332.2018.1444564>.

23. Goroshchuk O, Kolosenko I, Vidarsdottir L, Azimi A, Palm-Apergi C. Polo-like kinases and acute leukemia. *Oncogene*. 2019;38:1–16. <https://doi.org/10.1038/s41388-018-0443-5>.
24. Kim MS, Kim GM, Choi Y-J, Kim HJ, Kim Y-J, Jin W. TrkC promotes survival and growth of leukemia cells through Akt-mTOR-Dependent Up-Regulation of PLK-1 and Twist-1. *Mol Cells*. 2013;36:177–84. <https://doi.org/10.1007/s10059-0130061-6>
25. Kalbe B, Schlimm M, Wojcik S, Philippou S, Maberg D, Jansen F, et al. Olfactory signaling components and olfactory receptors are expressed in tubule cells of the human kidney. *Arch Biochem Biophys*. 2016;610:8-15. <https://doi.org/10.1016/j.abb.2016.09.017>.
26. Wang S, Chen B, Zhu Z, Zhang L, Zeng J, Xu G, et al. CDC20 overexpression leads to poor prognosis in solid tumors. *Medicine (Baltimore)*. 2018;97:e13832. <https://doi.org/10.1097/MD.0000000000013832>.
27. Simonetti G, Padella A, do Valle IF, Fontana MC, Fonzi E, Bruno S, et al. Aneuploid acute myeloid leukemia exhibits a signature of genomic alterations in the cell cycle and protein degradation machinery. *Cancer*. 2019;125:712–25. <https://doi.org/10.1002/cncr.31837>.
28. Salvador JM, Brown-Clay JD, Fornace AJ. Gadd45 in stress signaling, cell

cycle control, and apoptosis. *Adv Exp Med Biol.* 2013; 793:1–19.  
[https://doi.org/10.1007/978-1-4614-8289-5\\_1](https://doi.org/10.1007/978-1-4614-8289-5_1).

29. D'Angelo V, Crisci S, Casale F, Addeo R, Giuliano M, Pota E, et al. High Erk-1 activation and Gadd45a expression as prognostic markers in high risk pediatric haemolymphoproliferative diseases. *J Exp Clin Cancer Res.* 2009;28:39-46.  
<https://doi.org/10.1186/1756-9966-28-39>.

30. Vaisvilas M, Dirse V, Aleksuniene B, Tamuliene I, Cimbalistiene L, Utkus A, et al. Acute pre-B lymphoblastic leukemia and congenital anomalies in a child with a de novo 22q11.1q11.22 duplication. *Balk J Med Genet.* 2018;21:87–91.  
<https://doi.org/10.2478/bjmg-2018-0002>.

31. Xin Z, Soejima H, Higashimoto K, Yatsuki H, Zhu X, Satoh Y, et al. A novel imprinted gene, KCNQ1DN, within the WT2 critical region of human chromosome 11p15.5 and its reduced expression in Wilm's tumors. *J Biochim.* 2000;128:847-853. <https://doi.org/10.1093/oxfordjournals.jbchem.a022823>.

32. Yang F, Wu Q, Zhang L, Xie W, Sun X, Zhang Y, et al. The long noncoding RNA KCNQ1DN suppresses the survival of renal cell carcinoma cells through downregulating c-Myc. *J Cancer.* 2019;10:4662-70.  
<https://doi.org/10.7150/jca.29280>.

33. Ma X, Liu Y, Liu Y, Alexandrov L, Edmonson M, Gawad C, et al. Pan-cancer genome and transcriptome analyses of 1,699 pediatric leukemias and solid tumors. *Nature*. 2018;555:371-6. <https://doi.org/10.1038/nature25795>.
34. Oliveira ML, Akkapeddi P, Ribeiro D, Melão A, Barata JT. IL-7R-mediated signaling in T-cell acute lymphoblastic leukemia: An update. *Adv Biol Regul*. 2019;71:88–96.
35. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012;483:603–7. <https://doi.org/10.1038/nature11003>.
36. Yang J, Nie J, Ma X, Wei Y, Peng Y, Wei X. Targeting PI3K in cancer: mechanisms and advances in clinical trials. *Mol Cancer*. 2019;18:26-53. <https://doi.org/10.1186/s12943-019-0954-x>.
37. Sanchez V, Nichols C, Kim H, Gang E, Kim Y-M. Targeting PI3K signaling in acute lymphoblastic leukemia. *Int J Mol Sci*. 2019;20:412-25. <https://doi.org/10.3390/ijms20020412>.
38. Beer-Hammer S, Zebedin E, von Holleben M, Alferink J, Reis B, Dresing P, et al. The catalytic PI3K isoforms p110 $\gamma$  and p110 $\delta$  contribute to B cell development and maintenance, transformation, and proliferation. *J Leukoc Biol*. 2010;87:1083–95. <https://doi.org/10.1189/jlb.0809585>.

39. Montaña A, Forero-Castro M, Marchena-Mendoza D, Benito R, Hernández-Rivas J. New challenges in targeting signaling pathways in acute lymphoblastic leukemia by NGS approaches: An update. *Cancers*. 2018;10:1-24. <https://doi.org/10.3390/cancers10040110>.
40. Gyurina K, Kárai B, Ujfalusi A, Hevessy Z, Barna G, Jáksó P, et al. Coagulation FXIII A protein expression defines three novel sub-populations in pediatric B-cell progenitor acute lymphoblastic leukemia characterized by distinct gene expression signatures. *Front Oncol*. 2019;9:1063. <https://doi.org/10.3389/fonc.2019.01063>.
41. Vainchenker W and Constantinescu SN. JAK/STAT signaling in hematological malignancies. *Oncogene*. 2013;32:2601–13. <https://doi.org/10.1038/onc.2012.347>.
42. Minoguchi M, Minoguchi S, Aki D, Joo A, Yamamoto T, Yumioka T, et al. STAP-2/BKS, an adaptor/docking protein, modulates STAT3 activation in acute-phase response through its YXXQ motif. *J Biol Chem*. 2003;278:11182–9. <https://doi.org/10.1074/jbc.M211230200>.
43. Sekine Y. Adaptor protein STAP-2 modulates cellular signaling in immune systems. *Biol Pharm Bull*. 2014;37:185–94. <https://doi.org/10.1248/bpb.b13-00421>.
44. Ornitz DM, Itoh N. The fibroblast growth factor signaling pathway. *Wiley Interdiscip Rev Dev Biol*. 2015;4:215–66. <https://doi.org/10.1002/wdev.176>.

45. Su YZ, Wang CB, Zhou Y, Sun NT. Effects of changes in serum endostatin and fibroblast growth factor 19 on the chemotherapeutic sensitivity in acute myeloid leukemia patients. *Genet Mol Res.* 2015;14:5181-7. <https://doi.org/10.4238/2015>.
46. Dahl M, Kristensen L, Grønbaek K. Long non-coding RNAs guide the fine-tuning of gene regulation in B-cell development and malignancy. *Int J Mol Sci.* 2018;19:2475-500. <https://doi.org/10.3390/ijms19092475>.
47. James AR, Schroeder MP, Neumann M, Bastian L, Eckert C, Gökbuget N, et al. Long non-coding RNAs defining major subtypes of B cell precursor acute lymphoblastic leukemia. *J Hematol Oncol.* 2019;12:8-23. <https://doi.org/10.1186/s13045-018-0692-3>.
48. Wang W, Lyu C, Wang F, Wang C, Wu F, Li X, Gan S. Identification of potential signatures and their functions for acute lymphoblastic leukemia: A study based on the Cancer Genome Atlas. *Front. Genet.* 2021;12:656042. <https://doi.org/10.3389/fgene.2021.656042>.
49. Zaliova M, Kotrova M, Bresolin S, Stuchly J, Stary J, Hrusak O, et al. ETV6/RUNX1-like acute lymphoblastic leukemia: A novel B-cell precursor leukemia subtype associated with the CD27/CD44 immunophenotype. *Genes Chromosomes Cancer.* 2017;56:608–16. <https://doi.org/10.1002/gcc.22464>.
50. Edgar R. Gene expression omnibus: NCBI gene expression and hybridization

array data repository. Nucleic Acids Res. 2002;30:207–10.

<https://doi.org/10.1093/nar/30.1.207>.