



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO



FACULTAD DE ESTUDIOS SUPERIORES ZARAGOZA

*Análisis de los polimorfismos de un solo nucleótido en los genes CYP1A1 y CYP2E1 en niños con leucemia linfoblástica aguda*

TESIS

QUE PARA OBTENER EL TÍTULO DE:  
**QUÍMICO FARMACÉUTICO BIÓLOGO**

PRESENTA:

**ANAYELI MARTÍNEZ GARCÍA**  
311183991

Director de tesis:

**Dra. Silvia Jiménez Morales.**

Asesores:

**Dra. Soto Cruz María Isabel**  
**Dr. Mora Guevara José Luis Alfredo**

Ciudad de México (FES Zaragoza), Junio 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.



El presente trabajo se realizó bajo la dirección de la Dra. Silvia Jiménez Morales en el laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica (INMEGEN). El proyecto contó con el financiamiento del Consejo Nacional de Ciencia y Tecnología (CONACYT). Convocatoria Fronteras de la Ciencia (FOINS) 2016-01-2119.

## **AGRADECIMIENTOS Y DEDICATORIAS**

Me encuentro infinitamente agradecida con mi madre, que sin ella no hubiese sido posible culminar con éxito esta etapa de mi vida. Gracias por trabajar y esforzarte tanto en darme la oportunidad de cumplir mis sueños. No hay palabras suficientes para expresar mi gratitud y amor hacia ti. TE AMO.

No tengo palabras que puedan expresar lo agradecida que me encuentro con Lucy por haberme enseñado todo lo que necesite para llegar hasta donde me encuentro hoy, por que puso su confianza en mí y por todo el apoyo que me brindo desde el día que la conocí.

Agradezco a la Dra. Silvia Jiménez Morales por darme la oportunidad de tener mi primer acercamiento a la investigación, sus palabras de apoyo y por su atención brindada durante todo el proceso.

A mi casa de estudios la Universidad Nacional Autónoma de México, que me ha formado desde el nivel medio superior y hoy me permite concluir plenamente la licenciatura a través de la Facultad de Estudios Superiores Zaragoza.

Al Instituto Nacional de Medicina Genómica por el espacio brindado para desarrollar este trabajo.

A la comisión Nacional de Ciencia y Tecnología (CONACyT), por el apoyo financiero otorgado para este proyecto de investigación.

## INDICE

I. ABREVIATURAS .....	3
II. TABLAS .....	4
III. FIGURAS .....	5
IV. RESUMEN .....	6
1. INTRODUCCIÓN .....	7
2. MARCO TEÓRICO .....	9
2.1 Cáncer infantil.....	9
2.2 Leucemia linfoblástica aguda.....	9
2.2.1 Manifestaciones clínicas y diagnostico.....	10
2.2.2 Clasificación .....	11
2.2.3 Etiología .....	14
2.2.4 Citogenética .....	15
2.3 Polimorfismos .....	16
2.4 Estudios de asociación en la identificación de factores genéticos de riesgo para padecer LLA. ....	17
2.5 Genes que participan en el metabolismo de Xenobióticos y LLA. ....	19
2.5.1 Fase I .....	19
2.5.2 Fase II .....	22
3. PLANTEAMIENTO DEL PROBLEMA .....	24
4. PREGUNTA DE INVESTIGACIÓN .....	24
5. HIPÓTESIS .....	25
6. OBJETIVOS .....	25
6.1 Objetivo general.....	25
6.2 Objetivos particulares .....	25
7. MATERIAL Y METODOS.....	25

7.1 Características del estudio.....	25
7.2 Población de estudio .....	26
7.3 Estrategia general.....	27
7.4 Análisis molecular .....	28
7.4.2 Genotipificación por 5' exonucleasa .....	30
7.5 Análisis estadístico.....	33
8. RESULTADOS.....	34
8.1 Características clínicas y demográficas de la población analizada. ....	34
8.2 Análisis del rs4646903 en el gen <i>CYP1A1</i> . ....	35
8.3 Análisis del rs2031920 y el rs3813867 en el gen <i>CYP2E1</i> .....	36
8.4 Análisis de los SNPs en gen <i>NAT2</i> .....	37
8.5 Análisis del rs1800566 del gen <i>NQO1</i> .....	41
9. DISCUSIÓN .....	42
9.1 <i>CYP1A1</i> .....	43
9.2 <i>CYP2E1</i> .....	44
9.3 Gen <i>NQO1</i> .....	46
9.4 Gen <i>NAT2</i> .....	47
10. CONCLUSIONES .....	51
11. PERSPECTIVAS.....	51
12. REFERENCIAS.....	52

## **I. ABREVIATURAS**

ADN	Ácido desoxirribonucleico
CD	Clusters of differentiation
EDTA	Ácido etildiaminotetraacético
EGIL	The European Group for Immunological Classification of Acute Leukemia
FAB	Franco-Americano-Británico
OR	Odds ratio
OMS	Organización Mundial de la Salud
HWE	Equilibrio de Hardy-Weinberg (del inglés; Hardy-Weinberg equilibrium)
Hb	Hemoglobina
HLA-DR	Antígeno leucocitario humano clase II-DR
IC	Intervalo de confianza
LDH	Lactato deshidrogenasa
LLA	Leucemia linfoblástica aguda
MO	Medula ósea
PCR	Reacción en cadena de la polimerasa
SNP	Polimorfismo de un solo nucleótido
TDT	The transmission disequilibrium test
Xbs	Xenobiótico
NAT	N-acetiltransferasa
NQO1	Quinona Oxidoreductasa I
CYP	Citocromo

## **II. TABLAS**

Tabla 1. Características para la clasificación Franco-Américo-Británico de leucemia linfoblástica aguda.

Tabla 2. Clasificación inmunofenotípica de la leucemia linfoblástica

Tabla 3. Datos de las sondas TaqMan usadas para la discriminación alélica de los genes analizados en este trabajo.

Tabla 4. Características de la población analizada.

Tabla 5. Análisis de asociación entre el polimorfismo del gen CYP1A1 y leucemia linfoblástica aguda en pacientes mexicanos.

Tabla 6. Análisis de asociación entre los polimorfismos del gen CYP2E1 y leucemia linfoblástica aguda en pacientes mexicanos.

Tabla 7. Análisis de asociación de polimorfismos del gen NAT2 y leucemia linfoblástica aguda en niños mexicanos

Tabla 8. Análisis de asociación entre el polimorfismo del gen NQO1 y leucemia linfoblástica aguda en niños mexicanos.



### III. FIGURAS

Figura 1. Diagrama de hematopoyesis en donde se contempla la clasificación de leucemia de acuerdo a su linaje (mieloide o linfoide) y su evolución (agudas o crónicas).

Figura 2. Clasificación citológica de las leucemias linfoblásticas agudas. A) L1, B) L2 Y C) L3.

Figura 3. Condiciones de termociclador para PCR en tiempo final

Figura 4. Plot de discriminación alélica basada en la emisión de fluorescencia emitida por las sondas TaqMan.

#### IV. RESUMEN

La leucemia linfoblástica aguda es la neoplasia hematológica más frecuente en niños, se estima que esta enfermedad representa cerca del 80% de todas las leucemias presentes en niños con cáncer, la sobrevivencia en países desarrollados alcanza el 90%, sin embargo, para países en vías de desarrollo como México dicha cifra apenas alcanza el 60%. Se denota a la LLA como una enfermedad multifactorial, pues su etiología no ha sido del todo esclarecida.

En la actualidad se ha propuesto que polimorfismos en proteínas involucradas en el metabolismo de xenobióticos puede modificar el riesgo de desarrollar LLA, por ello el objetivo del presente trabajo fue conocer si los polimorfismos de los genes *CYP1A1*, *CYP2E1*, *NAT2* Y *NQO1*, Están asociados con la susceptibilidad de padecer leucemia linfoblástica aguda en niños mexicanos. Para ello se llevó a cabo un estudio de casos y controles, que incluyeron a 478 pacientes pediátricos diagnosticados con LLA como casos y a 284 individuos sanos como grupo control, todos de nacionalidad mexicana.

De los polimorfismos analizados se encontró asociación como variantes riesgo al rs1801280 (OR= 1.534; IC 95% 1.21-1.939, p=0.00034, rs1799929 (OR=1.963; IC 95% 1.549-2.487, p=1.740e-08) y al rs1208 (OR=1.46; IC95% 1.16-1.83, p=0.0012), en donde el riesgo se incrementa bajo un modelo recesivo (OR=2.20; IC95% 1.304-3.7110, OR=3.883; IC95%2.21-6.809 y OR=2.32; IC95%1.36-3.95, respectivamente), para el resto de los polimorfismos de *NAT2*, se encontró una probable asociación estadísticamente significativa como variantes de protección. En cuanto *CYP2E1* y *NQO1* no se encontró asociación estadísticamente significativa. Por lo cual se puede concluir que el gen *NAT2* se encuentra asociado con la leucemia linfoblástica aguda, incrementando el riesgo de desarrollar esta enfermedad en niños mexicanos.

## 1. INTRODUCCIÓN

Se estima que cada año se diagnostican en el mundo entre 175,000 y 250,000 nuevos casos de cáncer infantil, sin embargo, es probable que estas cifras estén subestimadas (1). De acuerdo a la organización mundial de la salud (OMS) el cáncer más frecuente en niños y adolescentes es la leucemia, con una incidencia de 2,127 casos nuevos hasta el 2018 y una mortalidad de 30%(2). En México, el cáncer infantil es un problema prioritario de la salud pública dado que las estimaciones señalan que hay de 2600 a 3120 casos nuevos de cáncer por año, de estos, las leucemias representan el 50.8%, de las cuales, el 80% son leucemia linfoblástica aguda (LLA)(3). Con respecto a la edad, 6.1% de los casos de cáncer infantil son menores de 1 año, 35.4% de 1 a 4 años, 26.5% de 5 a 9 años, 22.3% de 10 a 14 años y 9.9% de adolescentes entre 15 a 19 años (1).

En la actualidad la supervivencia libre de esta enfermedad alcanza hasta el 90% en países desarrollados, pues el constante desarrollo científico y tecnológico ha aumentado la posibilidad de un diagnóstico oportuno; aunado al diseño de esquemas terapéuticos adecuados al riesgo que tiene el paciente para recaer. El escenario en países en vías de desarrollo como México no es igual de favorable, puesto que solo se alcanza una sobrevida del 60% (1, 4, 5). Además de lo anterior aún existe un grupo de pacientes que no responde al tratamiento, lo que hace necesario la búsqueda de nuevas estrategias que permitan desarrollar esquemas de tratamiento más especializados para aumentar la sobrevida de los pacientes. Las herramientas genómicas prometen ser útiles para

generar conocimiento que impacte en el desarrollo de esquemas de diagnóstico, seguimiento y tratamiento a la medida del paciente (5)

Desde el punto de vista de la genómica se ha propuesto que polimorfismos en genes que codifican a proteínas involucradas en el metabolismo de sustancias exógenas como contaminantes, pesticidas, alcohol e incluso medicamentos, tienen un papel en el desarrollo de diferentes neoplasias como las leucemias, es por ello que el presente trabajo tuvo como objetivo conocer si existe asociación entre los polimorfismos de un solo nucleótido (6) presentes en los genes CYP1A1, CYP2E1, NAT2 y NQO1 y la LLA.

## **2. MARCO TEÓRICO**

### **2.1 Cáncer infantil**

En el mundo cerca de la mitad de los casos reportados de cáncer infantil son hematológicos, de los cuales las leucemias ocupan el primer lugar. Leucemia es el término que se usa para definir un grupo de enfermedades malignas en la sangre y se caracteriza por la proliferación clonal y descontrolada de cualquier célula del tejido hematopoyético. Las células inmaduras y transformadas ocupan la médula ósea y con el avance de la enfermedad, migran a la circulación periférica infiltrándose a otros tejidos (7).

### **2.2 Leucemia linfoblástica aguda**

La LLA es la neoplasia hematológica de la estirpe linfoide que deriva de la transformación maligna de una célula progenitora capaz de generar células idénticas con diversas anomalías que les confieren resistencia a la apoptosis y desregulación del ciclo celular(8).

De acuerdo al GLOBOCAN en México se diagnostican anualmente 1470 casos de cáncer en niños de 0 a 14 años, en donde las leucemias son las neoplasias más frecuentes (28.9%), de los cuales, la LLA constituye hasta el 80% de los casos (2). En países desarrollados se alcanza una tasa de curación de hasta el 90%, sin embargo, en países en vías de desarrollo como México, la supervivencia es menor al 60% a 5 años (1).

### **2.2.1 Manifestaciones clínicas y diagnóstico**

Las manifestaciones clínicas de la LLA suelen presentarse de manera gradual o abrupta cuando disminuye la hematopoyesis normal, a consecuencia de la ocupación de medula ósea. Por lo general el paciente presenta: palidez, fatiga, taquicardia, disnea, petequias, purpura, hemorragias, fiebre y cuadros infecciosos con mayor frecuencia. Cuando la enfermedad avanza y hay infiltración extramedular, el paciente presenta: dolor óseo, esplenomegalia, hepatomegalia, adenopatías y en casos muy avanzados infiltración a sistema nervioso central, hígado, bazo, pleura y testículo (8, 9).

En los estudios de laboratorio, hasta el 60 % de los pacientes muestran recuentos celulares por debajo de los valores de referencia ( $Hb < 8g/dl$ , plaquetas  $< 50 \times 10^9/L$  y leucocitos  $< 10 \times 10^9/L$ ); sin embargo, la cantidad de células en sangre periférica dependerá de la proliferación existente en la medula ósea y del tiempo de evolución de la enfermedad(9, 10). El aspirado de médula ósea (MO) permite observar al porcentaje de ocupación por parte de las células neoplásicas, que en casos avanzados puede ser hasta del 100%. Usualmente se sospecha de una leucemia cuando la cantidad de blastos en MO es mayor al 5% y es confirmatoria cuando alcanzan el 20%(8, 11).

Otros estudios complementan el diagnóstico de LLA haciendo evidente el incremento del metabolismo celular el cual es proporcional al desarrollo del tumor. La lactato deshidrogenasa (LDH) se ve aumentada por el proceso constante de destrucción celular, se observa hipercalcemia como resultado de la resorción de hueso en MO y uricemia por la tasa de recambio celular. Es posible ver resultados

anormales en otros estudios que evalúan el funcionamiento de órganos debido al daño producido por la infiltración celular, tal es el caso de hígado y riñón. En sangre periférica se suele observar una cantidad considerable de blastos, parámetro básico para el diagnóstico de leucemia, sin embargo, en la mayoría de los casos la baja cantidad de células dificulta el diagnóstico usando sólo este parámetro(8).

### 2.2.2 Clasificación

Las leucemias se dividen de acuerdo a su célula progenitora como mieloblásticas o linfoblásticas y con respecto a su evolución en agudas o crónicas. La evolución aguda de la enfermedad se caracteriza por un desarrollo rápido y la presencia de precursores inmaduros. En cuanto a la evolución crónica, el desarrollo de la enfermedad es lento y las células presentes están en etapas de maduración más avanzadas(4, 8). La clasificación de esta enfermedad se hace con base en la morfología celular, propiedades citoquímicas e inmunofenotipo.

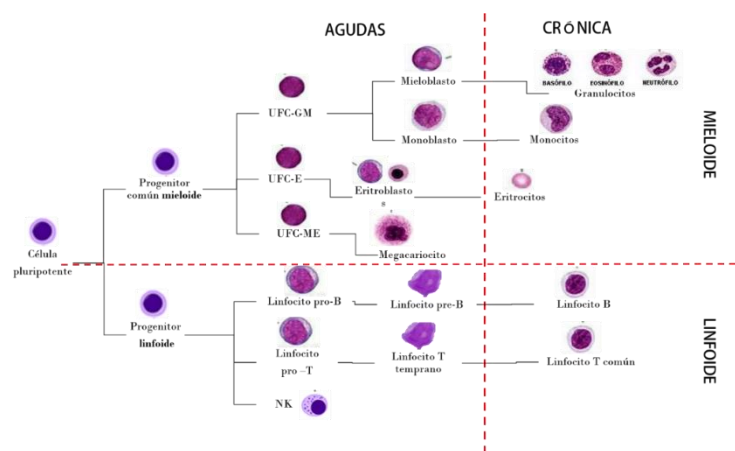


Figura 1. Diagrama de hematopoyesis normal y leucémica. Se esquematiza clasificación de leucemia de acuerdo a su linaje (mieloide o linfóide) y su evolución (agudas o crónicas).

La leucemia de origen linfóide se distingue por la presencia de blastos pequeños de núcleo grande con cromatina homogénea y escaso citoplasma; son negativas para mieloperoxidasa y negro sudan, su inmunofenotipo se representa por la positividad a los marcadores: CD2, CD10, CD3, CD4, CD7, CD19 y CD20.

La primera subclasificación sugerida para LLA se realizó por el grupo Francés-Americano-Británico (12), el cual, tomando en cuenta las características morfológicas celulares, divide a la LLA en los subtipos L1, L2 y L3 (Tabla1, Figura2). Sin embargo, en la actualidad, el diagnóstico y la clasificación estándar de esta enfermedad se basa en el inmunofenotipo(13).

Tabla1. Clasificación Franco-Americano-Británico (12) de leucemia linfoblástica aguda.

<b>Característica citológica</b>	<b>L1</b>	<b>L2</b>	<b>L3</b>
<b>Tamaño celular</b>	Predominancia de células pequeñas	Grandes o con tamaño heterogéneo	Grandes y heterogéneas
<b>Cromatina</b>	Homogénea y fina	Variable	Punteada, fina y homogénea
<b>Núcleo</b>	Regular, ocasionalmente hendido.	Irregular, comúnmente con hendidura	Regular, oval o redondo.
<b>nucléolos</b>	Ausentes	Uno a dos	Uno a dos prominentes
<b>Citoplasma</b>	Escaso, ligeramente basófilo.	Abundante y basófilo	Abundante y basófilo



Tomada y modificada de Carrol 2016, (11)

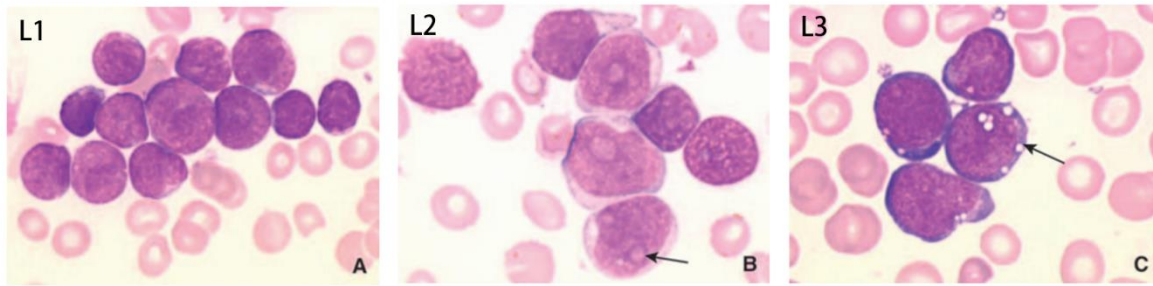


Figura 2. Clasificación citológica de las leucemias linfoblásticas agudas. A) L1, B) L2 Y C) L3. Tomado y modificado de Michel G, 2008. (14)

El inmunofenotipo se determina con citometría de flujo y anticuerpos monoclonales específicos para diversos antígenos expresados por las células y que difieren según el tipo de célula y su etapa de maduración (Tabla 2.).

Tabla 2. Clasificación de leucemia linfoblástica con respecto a marcadores de superficie.

Subtipo	Inmunofenotipo	Frecuencia	Pronostico
<b>Pro-B</b>	HLA-DR, TdT, CD19,CD79a,CD22, CD10 <sup>v</sup>	5%	Desfavorable
<b>Pre-B temprana</b>	HLA-DR, TdT, CD19,CD79A,CD22, CD10+, cIgM	63%	CD10+ buen pronostico
<b>Pre- B</b>	TdT, HLA-DR, CD19, CD20,IgC-μ, IgS- μ.	16%	Buen
<b>Célula B</b>	HLA-DR, CD19,CD20,CR2, μ,κ o λ	15%	Desfavorable
<b>T-temprana</b>	TdT, CD7,CD5,CD3c, CD2 <sup>v</sup>	10%	Desfavorable

<b>T-cortical</b>	TdT, CD7, CD2, CD4 y/o CD8,CD5,CD1	1	Desfavorable
<b>T-madura</b>	CD3, CD7,CD5, CD4 o CD8	2	Desfavorable

Tomada de Arriaga,2019 (13)

Existen casos reportados de leucemias agudas que expresan características de más de un linaje hematopoyético y se han definido como leucemias de linaje mixto o leucemias bifenotípicas, la incidencia no es mayor al 5% y actualmente cuentan con una clasificación dada por EGIL (The European Group for Immunological Classification of Acute Leukemia) (15).

### **2.2.3 Etiología**

La LLA es una entidad compleja, en donde los agentes etiológicos que influyen en su desarrollo son tanto ambientales como genéticos.

Dentro de los factores ambientales se encuentra: la exposición a radiaciones ionizantes, el consumo de tabaco, alcohol, café y medicamentos con efecto leucomogénico (ejemplo. alquilantes, inhibidores de topoisomerasa II), la exposición a carcinogénicos (hidrocarburos como el benceno) y pesticidas; así como, infecciones. Éstas últimas, se ha sugerido que están asociados con LLA, ya que en la edad pediátrica el sistema inmune no se encuentra del todo desarrollado, por lo cual hay una mayor susceptibilidad a los efectos oncogénicos de determinados agentes patógenos como los virus, Epstein Barr y los HTLV I y II; aunque hasta el momento, no han presentado una clara asociación (4, 11, 16).

Dentro de los factores genéticos de riesgo se encuentran, el síndrome de Down, anemia de Fanconi, síndrome de Bloom y ataxia telangictasica, ya que tienen diversos mecanismos que promueven el desarrollo de una neoplasia como; la fragilidad cromosómica, mayor susceptibilidad a la replicación viral y alteraciones en diversos genes (16).

Desde hace unos años se ha reportado que la combinación de factores ambientales y la presencia de mutaciones o alteraciones a nivel génico, contribuyen al desarrollo de diversas patologías complejas como el cáncer. Para la LLA, se han asociado varios polimorfismos de un solo nucleótido (SNP por sus siglas en inglés) en genes que codifican para enzimas que participan en el metabolismo de xenobióticos con el riesgo de desarrollar la enfermedad, presentar recaída temprana, toxicidad a quimioterapia, menor resistencia a agentes antineoplásicos y menor supervivencia (5, 17, 18).

#### **2.2.4 Citogenética**

El amplio estudio de la LLA desde la perspectiva de los análisis génicos ha revelado que alteraciones a nivel de ácido desoxirribonucleico (19) están asociadas al desarrollo de esa neoplasia y se ha reportado que dichos cambios están presentes hasta en el 75% de los pacientes (20). Dentro de las anomalías identificadas están:

- Aneuploidías o alteración en el número de cromosomas. Las hiperdiploidias (>50 cromosomas) están asociadas a un excelente pronóstico y se presenta en 30%

de los casos; mientras que las hipodiploidias (< 44 cromosomas) son de muy mal pronóstico y se presentan en 3% de los casos (21).

- Translocaciones (cambio de segmentos entre dos cromosomas), con respecto LLA, son las más estudiadas, pues por lo general involucran genes relacionados con los factores de transcripción hematopoyéticos, moduladores epigenéticos, proteínas constitutivas, entre otros. Para el linaje de células B, las translocaciones de valor pronóstico son: *t(12;21) (TEL-AML/ETV6-RUNX1)*, *t(1;19) (E2A-PBX1/TCF3-PBX1)* y *t(9;22) (BCR-ABL1)* y fusiones que involucran al gen *MLL (MLL-AF4)*, principalmente). En cuanto al linaje de células T, las anomalías de interés son: mutaciones de *NOTCH1* y los rearrreglos en *TLX1-HOX11*, *TLX3-HOX11L2*, *LYLI*, *TALI* y *MLL(5, 21)*.
- Mutaciones puntuales en genes que involucran proteínas con roles clave en el desarrollo linfóide (*PAX5, IKZF1, EBF1, LMO2*), en la regulación del ciclo celular como genes supresores de tumor (*CDKN2A/CDKN2B, PTEN, RB1, TP53*), y aquellos que participan como señalizadores linfoides (*BTLA, CD200*), entre otros (20).

### **2.3 Polimorfismos**

Los polimorfismos constituyen un elemento clave en la investigación y la práctica de la genética humana. Son variaciones en la secuencia del genoma que se presentan en más del 1% de la población abierta (22).

Después de conocer la secuencia del ADN, se determinó que solo el 99% de la identidad de la secuencia será compartida entre los individuos, el resto se conoce

como variabilidad genética, la cual se atribuye a los diferentes polimorfismos (inserciones, deleciones, cambios de numero de secuencias repetidas, etc) y principalmente a los SNP (del inglés, *single nucleotide polymorphisms*). Los SNPs son los polimorfismos más frecuentes, en su mayoría se presentan de forma bialélica y en función de su localización dentro del gen se pueden clasificar en codificantes y no codificantes. A su vez, los SNP codificantes pueden ser sinónimos o no sinónimos, si no modifican el aminoácido para el que codifican o si lo cambian, respectivamente. Estos últimos pueden alterar la función de la proteína (23).

Se sabe que algunas de estas variantes confieren alto riesgo para padecer la enfermedad, modifican su gravedad o se asocian a la respuesta al tratamiento. En este último caso, existen algunas variantes que requieren de su genotipificación previo al tratamiento (22).

#### **2.4 Estudios de asociación en la identificación de factores genéticos de riesgo para padecer LLA.**

Una de las estrategias que ha llevado a la identificación de SNPs potencialmente relevantes en el tratamiento de una enfermedad han sido los estudios de asociación. Estos estudios permiten conocer si hay relación estadística entre un marcador y un rasgo o enfermedad. Es una herramienta que se ha usado por diversos investigadores para conocer la contribución genética en enfermedades complejas como la diabetes, hipertensión, cáncer, etc. Los marcadores pueden ser genes, regiones de cromosomas o polimorfismos (24, 25). Este tipo de estudios se basa en el hallazgo de que si un el alelo de un locus se encuentra diferencialmente distribuido entre los casos y los controles (ya sea en mayor o menor frecuencia),

entonces el marcador se encuentra asociado a la enfermedad(22). Los diseños de estudios de asociación más comunes son:

- a) Casos y controles: consiste en la comparación entre individuos no relacionados, los cuales son clasificados en dos grupos; casos (individuos que presentan la enfermedad o el rasgo de interés) y controles (aquellos que no presentan la enfermedad). Este diseño suele ser ideal para el estudio de enfermedades raras o complejas, además de ser relativamente más barato y fácil de realizar(26).
- b) Cohorte: en este tipo de diseño se selecciona a un grupo de individuos con características definidas a las cuales se les dará un seguimiento, el tiempo aproximado en el que se espera la aparición de la enfermedad o el rasgo de interés. Este estudio por lo general es ideal para la evaluación de incidencia o de historia de la enfermedad; sin embargo, son caros y requieren un periodo de seguimiento en donde puede existir la pérdida de los integrantes(27).
- c) Basados en relación familiar: el más simple de los diseños de este tipo es aplicar la prueba de desequilibrio de transmisión o TDT (del inglés, the transmission disequilibrium test). Consiste en el análisis de tríos, la persona afectada y sus dos padres. En general este diseño evita problemas de estratificación en la población (sesgos en la selección de la muestra), sin embargo, muchas veces es difícil aplicarlo porque se requiere de muchas familias(28).

## **2.5 Genes que participan en el metabolismo de Xenobióticos y LLA.**

Los humanos se encuentran en constante exposición a todo tipo de sustancias que pueden resultar dañinas para la salud, entre ellas está: químicos ambientales, aditivos de alimentos, agentes contaminantes y medicamentos; este tipo de sustancias exógenas se les denomina Xenobióticos(Xbs) (29).

Los xenobióticos son sustancias que no son parte de la composición habitual del cuerpo, pero que son capaces de acceder a su interior (30). Los humanos tienen un sistema de defensa contra estas sustancias exógenas en donde se transforman por un conjunto de procesos enzimáticos cuyo objetivo es incrementar su hidrofilia para facilitar su neutralización y eliminación; dicha transformación se lleva a cabo en dos etapas:

### **2.5.1 Fase I**

Consta de reacciones del tipo oxidación, hidrolisis y reducción, cuya finalidad es formar moléculas más polares que faciliten su excreción. Durante esta etapa se ven involucradas varias enzimas, siendo el complejo citocromo p450 (CYP450) el más importante de éstas por la gran cantidad de reacciones en las que están involucrados.

#### **2.5.1.1 Citocromo 450.**

Los CYP450 o P450, es un conjunto de proteínas monoxigenasas que contienen un grupo hemo, se caracterizan por absorber la luz a 450nm en su estado reducido y en presencia de monóxido de carbono (29, 30). Estas proteínas se encuentran

distribuidas en diferentes tejidos entre ellos: la piel, pulmón, cerebro, tejido hematopoyético y de forma más abundante en hígado e intestino. Esta superfamilia de proteínas constan con un sistema de nomenclatura donde se usa el prefijo “CYP” seguido del número de la familia, la identidad de la subfamilia y el número de miembro de esta(31).

A través de los años se ha estudiado la variabilidad genética de P450 y su relación con la capacidad de metabolizar diferentes sustancias, así como su relación con respecto a algunas enfermedades como el cáncer. Dentro de estos estudios se demostró que la presencia de algunos SNPs en estos genes se encuentran relacionados con la actividad enzimática ya sea que resulte normal, disminuida, aumentada o ausente(29). *CYP1A1* Y *CYP2E1* están entre los genes más estudiados en LLA.

#### 2.5.1.1.1 CYP1A1

Este gen se localiza en el cromosoma 15 en la sección q24.1, tiene 6 intrones y 7 exones cuyo tamaño es de 6,069 bases. *CYP1A1* codifica para la hemoproteína de la familia 1, subfamilia A, miembro 1 y tiene como sustratos a los hidrocarburos aromáticos poli-cíclicos, benzopirenos, dioxinas halogenadas, estrógenos, entre otros. Se expresa en pulmón, hígado, cerebro, corazón y linfocitos.(32) Es un gen en donde se han encontrado múltiples SNPs, entre los cuales está la variante *CYP1A1\*2* (33), la cual consiste en un cambio de timina (T) por una citosina (C) en la posición 3801, el rs4646903 se encuentra en la región 3´ de la zona no codificante y a pesar de no modificar la enzima, se ha encontrado un aumento en la expresión y estabilidad del mRNA de *CYP1A1*, incrementando de este modo su



actividad (34-36). Dado su rol en el metabolismo de Xbs, estudios previos se han interesado en su rol en enfermedades como el cáncer, con respecto a LLA se ha asociado al rs4646903 con el aumento del riesgo de esta enfermedad en población hispana de forma contradictoria a lo encontrado por Lu y colaboradores, en donde no reportaron asociación con Caucásicos o Asiáticos, estos resultados inconclusos también se han observado en población de la india (37)(36, 38, 39).

#### 2.5.1.1.2 CYP2E1

Este gen se encuentra ubicado en el cromosoma 10 en la sección q26.3, tiene un tamaño de 40,815 bases, 8 intrones y 9 exones. Codifica para la hemoproteína de la familia 2, subfamilia E, miembro 1 y tiene como sustratos la clorzoxazona, N,N-dimetilalanilina, trimetidiona, acetaminofén, halotano, alcohol, benceno, acetona, nitrosamina entre otros. Además de la piel, pulmón, cerebro, tejido hematopoyético, hígado e intestino, *CYP2E1* también se expresa en MO y el endotelio(32). El estudio de las variantes génicas permitió encontrar dos polimorfismos en la región promotora 5' el rs2031920 (RsaI) y el rs3813867 (PstI) (40). Ambos polimorfismos se han asociado con diversas enfermedades incluyendo cáncer de cavidad oral, nasofaringe, estómago y esófago (41). Su relación con LLA ha sido poco explorada (42). Se ha descrito que los alelos *CYP2E1*\* 5B y 6\* incrementan hasta tres veces más el riesgo a padecer LLA, en algunos de los estudios dicha asociación se encuentra considerando que algunos de los progenitores presentan condiciones tales como alcoholismo (17, 43, 44).

## 2.5.2 Fase II

En esta etapa se involucran enzimas denominadas transferasas actúan sobre los Xbs o los metabolitos resultantes de la fase I y dan paso a reacciones de conjugación con sustancias endógenas que facilitan su transporte por el organismo y su rápida excreción por medio de la orina o la bilis(45). Dentro de este grupo se encuentran la N-Acetiltransferasa 2 (NAT2) y NAD(P)H: quinona oxidoreductasa 1 (NQ01)

### 2.5.2.1 NAT2

Este gen se encuentra en el brazo corto del cromosoma 8, está conformado por 14,635 bases, se compone por dos exones y un intrón. Codifica para la proteína N-Acetiltransferasa la cual forma parte de la Fase II de la biotransformación de Xbs y cataliza la acetilación usando como cofactor a la Acetil-CoA. Dentro de sus sustratos están las amidas aromáticas o compuestos con hidracinas (sulfametazina, isoniazida, hidralazina y sulfamidas)(31).

*NAT2* es altamente polimórfico y es por ello que hay un consenso en su nomenclatura con respecto los fenotipos de acetilación ya sea rápida, lenta o intermedia (<http://nat.mbg.duth.gr/>). Algunos de los polimorfismos que se encuentran en este gen son: rs1801280 (Ile114Thr), rs1799930 (Arg197Gln), rs1799931 (Lys286Glu), rs1208 (Lys268Arg), rs1041983 y rs1799929; Las variaciones de este gen se han asociado con la modificación de la actividad de la enzima o con su estabilidad. es por ello que ha sido de gran interés para los investigadores entender su relación con el desarrollo de cáncer (18, 46).

### 2.5.2.2 NQO1

El gen que codifica a NAD(P)H: quinona oxidoreductasa 1, se encuentra ubicado en el cromosoma 16q22.1, consta de 19,673 bases y está conformado por 5 intrones y 6 exones. La enzima producto forma parte de la segunda línea de defensa, cuya propiedad es la transferencia de dos electrones usando NADH o NADPH como agente reductor, algunos de sus sustratos son: quinonas exógenas y endógenas, diclofenolindolfenol, compuestos con grupos nitro y azo y azul de metileno (47). Este gen tiene más de 93 SNPs, entre ellos el más investigado es el C609T (rs1800566), el cual consiste en un cambio de una citosina por una timina en el nucleótido 609. Este SNP provoca un cambio de aminoácidos en la posición 187 de prolina a serina (Pro187Ser), modificación asociada con la pérdida de la función de la enzima(48). La actividad de NQO1 ayuda a la protección de la célula frente a radicales libres y metabolitos tóxicos generados por la reducción catalizada por CYP450 u otras enzimas. Se ha reportado que la variante rs1800566 está asociada como variante de riesgo para las leucemias agudas, sin embargo, para la LLA, los estudios han reportado resultados controversiales. Por un lado asociación con esta enfermedad bajo un estudio familiar y en otros en donde no hay asociación(49, 50) (51).

### **3. PLANTEAMIENTO DEL PROBLEMA**

En México la LLA es un problema de salud pública, es la neoplasia hematológica más común en edad pediátrica y ocupa entre la segunda y tercera posición dentro de todos los casos de cáncer en niños de 0 a 14 años. La LLA ha sido identificado como una enfermedad multifactorial, sin embargo, aún se desconocen muchos de los factores de riesgo genéticos implicados en su desarrollo.

Se ha propuesto que la interacción ambiental y la susceptibilidad genética son un factor de riesgo para desarrollar cáncer, El complejo enzimático Citocromo P450, la N-Acetiltransferasa y la NAD(p) quinona oxidoreductasa I, son enzimas involucradas en la biotransformación de sustancias exógenas, por esa razón investigadores han estudiado SNPs presentes en sus genes y su relación con la susceptibilidad de desarrollar diferentes tipos de cáncer, entre ellos la LLA. En la Actualidad los estudios que asocian a los SNPs en genes del metabolismo de Xbs y la LLA son muy pocos, por lo tanto, en esta tesis se pretende conocer la asociación existente entre los polimorfismos de los genes *CYP1A1*, *CYP2E1*, *NAT2* y *NQO1* y la susceptibilidad a desarrollar LLA en niños mexicanos.

### **4. PREGUNTA DE INVESTIGACIÓN**

¿Existe asociación entre los polimorfismos de los genes *CYP1A*, *CYP2E1*, *NAT2* y *NQO1* y la susceptibilidad a desarrollar leucemia linfoblástica aguda en niños mexicanos?

## **5. HIPÓTESIS**

Si SNPs en genes *CYP1A1*, *CYP2E*, *NAT2* y *NQO1* son más frecuente en niños con LLA que en niños sin la enfermedad, entonces estos SNPs se asociarán con un incremento en el riesgo para padecer la enfermedad en niños mexicanos

## **6. OBJETIVOS**

### **6.1 Objetivo general**

Conocer si los SNPs en los genes *CYP1A1*, *CYP2E1*, *NAT2* y *NAO1* se asocian a la susceptibilidad de padecer leucemia linfoblástica aguda en niños mexicanos.

### **6.2 Objetivos particulares**

Conocer la frecuencia de SNPs en los genes *CYP1A1*, *CYP2E1*, *NAT2* y *NQO1* en niños mexicanos.

Determinar si los SNPs en estos genes confieren riesgo o protección para el desarrollo de LLA en niños mexicanos

Determinar si existe asociación entre los SNPs evaluados y variables clínicas de la enfermedad.

## **7. MATERIAL Y METODOS**

### **7.1 Características del estudio.**

- Diseño: Estudio de asociación casos y controles,
- Tipo: observacional, prospectivo y comparativo

## **7.2 Población de estudio**

### **Casos**

#### Criterios de inclusión

- Edad menor a 18 años
- Diagnóstico clínico e histopatológico de LLA de *novo*
- Sin tratamiento leucémico previo
- Consentimiento informado firmado por padres o tutores

#### Criterios de exclusión

- Presencia de síndromes genéticos
- Presencia de otras neoplasias
- Enfermedades hematológicas o enfermedades infectocontagiosas

#### Criterios de eliminación

- Pacientes cuyos padres decidieron salirse del estudio,
- Muestras de ADN sin la calidad adecuada para las determinaciones moleculares
- Pacientes que durante el seguimiento no continuaron con el tratamiento.

### **Controles**

#### Criterios de inclusión

- Edad menor a 18 años
- Sin LLA o enfermedades hematológicas
- Consentimiento informado firmado por padres o tutores

#### Criterios de exclusión

- Presencia de síndromes genéticos
- Presencia de otras neoplasias
- Enfermedades infectocontagiosas

#### Criterios de eliminación

- Pacientes cuyos padres decidieron salirse del estudio,
- Muestras de ADN sin la calidad adecuada para las determinaciones moleculares

### **7.3 Estrategia general**

Para lograr los objetivos planteados se llevó a cabo la selección de los individuos pertenecientes a los grupos de casos y controles de acuerdo a los criterios mencionados previamente, después de firmar el consentimiento informado por los tutores y con ayuda del personal de salud, se obtuvieron muestras de sangre o saliva de cada uno de ellos. Una vez obtenida la muestra se llevó a cabo la extracción de ADN, cuya calidad fue evaluada a través de su integridad. En lo que respecta al análisis molecular para las variantes de interés, las muestras fueron genotipificadas a través de la técnica de 5' exonucleasa usando sondas TaqMan, finalmente, el análisis estadístico se llevó a cabo mediante una prueba de Chi cuadrada para conocer los datos de asociación (Figura 3).

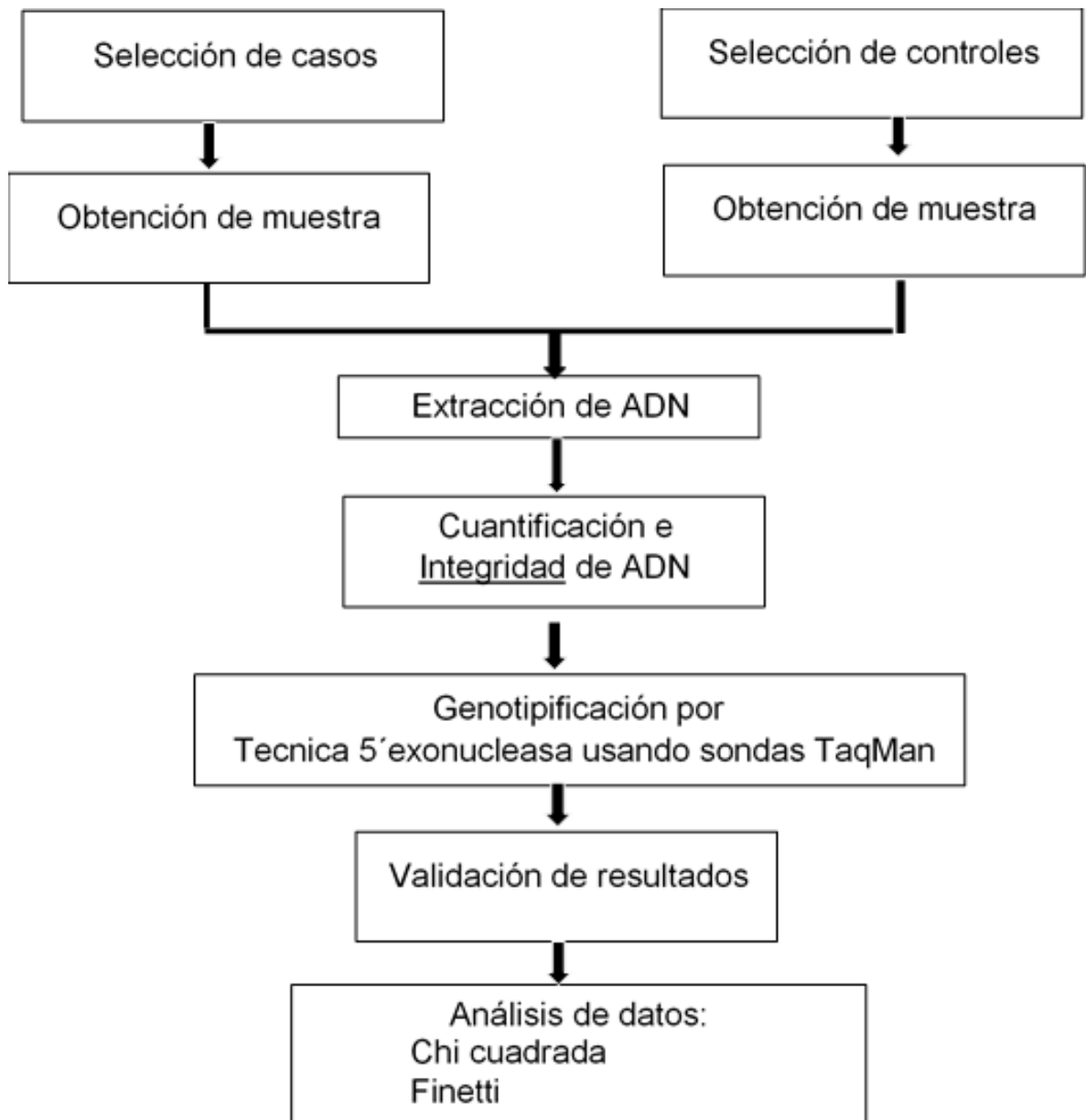


Figura 3. Estrategia general.

## 7.4 Análisis molecular

### 7.4.1. Extracción de ADN

La extracción del ADN fue a partir de muestras de sangre periférica y de saliva. La muestra de sangre periférica fue recolectada en tubos vacoutainer con EDTA y se



utilizó el kit de QIAGEN Genra puregene Blood siguiendo las especificaciones del proveedor.

Después de recolectar las muestras, estas se centrifugaron a 3,500 rpm durante 15 min; posteriormente se separó la capa de leucocitos y se transfirieron a un tubo Falcón de 15 mL. Al botón de células se le agregaron 6mL de solución de lisis de eritrocitos y se agitó fuertemente con el vortex durante 30s. En seguida se centrifugó nuevamente a 3000 rpm durante 10 min, en donde se decantó el sobrenadante para obtener el pellet de leucocitos. Se hizo la extracción de ADN genómico, adicionando 2mL de buffer de lisis de leucocitos agitando vigorosamente en vortex durante 1 min y se incubó durante 24h.

Después de este tiempo, se adicionó 1ml de solución para precipitar proteínas seguido de agitación constante en vortex por 1 min. Se centrifugó nuevamente a 3000 rpm, 10 min y cuidadosamente se transfirió el sobrenadante a otro tubo falcón de 15 mL. A este tubo, se le agregaron 5 mL de isopropanol frío para precipitar el ADN. El ADN se separó de la solución después de centrifugar durante 10 min a 3000 rpm, después de dos lavados con etanol al 75%, el ADN se diluyó en 200 µL de Buffer TE.

Para muestras de saliva, se usó el kit ORAGENE bajo las especificaciones del proveedor. El primer paso fue la incubación de la muestra a 50°C por 4h, después se transfirieron 500µL de muestra a un tubo Eppendorf de 2 ml, al cual se le adicionaron 20µL de solución para precipitar proteínas prepIT-L2P (DNA genotek). Se agitó brevemente en vortex y se dejó incubando en frío durante 10 min. Posteriormente, se centrifugó el tubo a 12,000 rpm durante 10 min, se recuperó

cuidadosamente el sobrenadante en un tubo limpio, al cual se le adicionaron 600µL de etanol absoluto y se incubó a temperatura ambiente por 30 min. En seguida, se centrifugó el tubo a 12,000 rpm por 2 min, y se lavó la pastilla de ADN con etanol al 75%. La hidratación del DNA se realizó con 150 µL Buffer TE.

#### 7.4.2 Genotipificación por 5' exonucleasa

Los SNP en los genes *CYP1A1*, *CYP2E1*, *NAT2* Y *NQO1* se analizaron mediante la técnica 5' exonucleasa usando sondas TaqMan (Applied Biosystem Foster City, CA, USA,) (Tabla 3).

Tabla 3. Datos de las sondas TaqMan usadas para la discriminación alélica de los genes *CYP1A1*, *CYP2E1*, *NAT2* y *NQO1*, analizados en este trabajo.

GEN	POLIMORFIMOS	CLAVE	ALELO VIC/FAM
<b><i>CYP1A1</i></b>	rs4646903		C/T
<b><i>CYP2E1</i></b>	rs2031920		C/T
	rs3813867	C_2431875_10	G/C
<b><i>NAT2</i></b>	rs1041983	C_8684085_20	C/T
	rs1801280	C_1204093_20	C/T
	rs1799929	C_1204092_20	C/T
	rs1799930	C_1204091_10	A/G
	rs1208	C_572769_20	A/G
	rs1799931	C_572770_20	A/G
<b><i>NQO1</i></b>	rs1800566	C_2091255_30	C/T

#### 7.4.2.1 Condiciones de la reacción en cadena de la polimerasa.

Se llevó a cabo la reacción de cadena de la polimerasa (PCR) a un volumen final de 5  $\mu$ L, para lo cual se usaron 2.5  $\mu$ l de master mix (dNTPs, TaqGold, MgCl<sub>2</sub>), 0.08  $\mu$ L de sonda y 2.4  $\mu$ L de agua grado molecular por cada muestra. La reacción de PCR se realizó en un termociclador GeneAmp PCR system 9700, bajo el siguiente programa: un ciclo a 50°C por 2 min, un ciclo a 95° por 10 min, 45 ciclos de 95° por 15 s y 60°C por 1 min y un tiempo indefinido a 4° para almacenamiento (Figura 4).

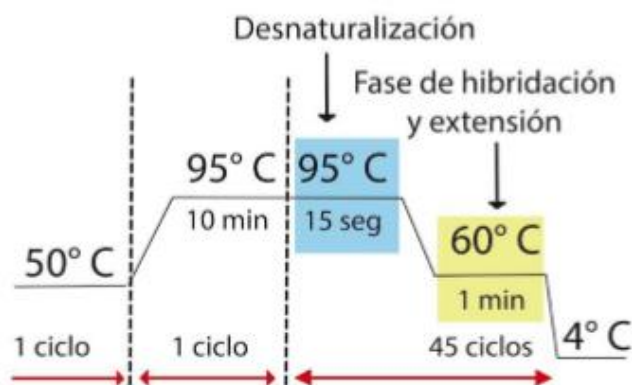


Figura 4. Condiciones de termociclador para PCR en

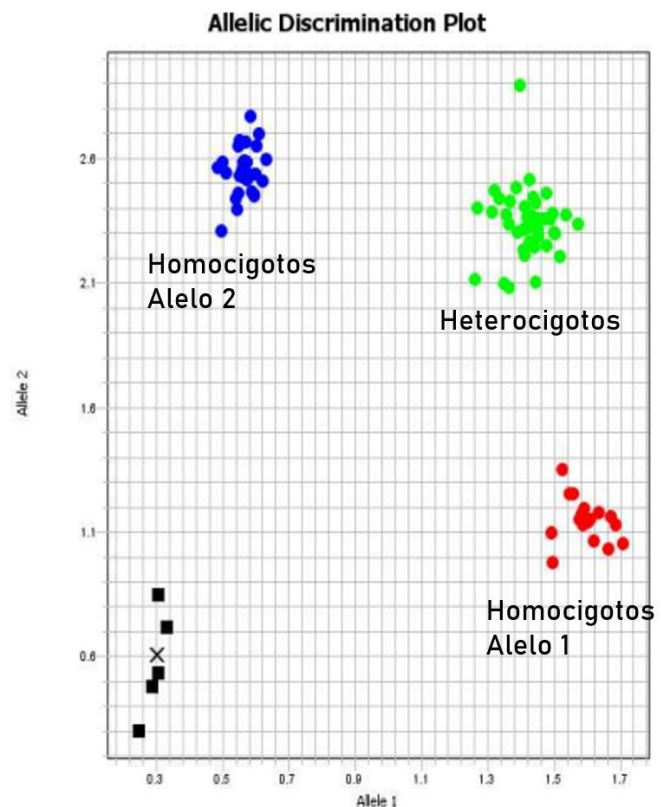
#### 7.4.2.2 Técnica de la 5' exonucleasa

Para cada ensayo de discriminación alélica se diseñan 2 sondas específicas marcadas en el extremo 5' con fluorocromos diferentes, VIC para alelo 1 y FAM para el alelo 2. Cada sonda contiene en el extremo 3' un "quencher" el cual inhibe la emisión de fluorescencia de la sonda mientras está intacta. Durante la reacción de PCR, los primers y las sondas hibridan con su secuencia complementaria, si contiene la secuencia polimórfica la AmpliTaq Gold, que tiene actividad de ADN

polimerasa y de exonucleasa 5'-3', va a digerir la sonda durante el proceso de amplificación liberando al fluorocromo de la acción del “quencher” y así permitiendo la emisión de fluorescencia.

Al finalizar la PCR, se cuantifica la fluorescencia obtenida en el equipo de tiempo real QuantStudio5 Real-Time PCR system. La diferenciación de los alelos se determina con el tipo de fluorescencia emitida (VIC o FAM), tal como se muestra en la Figura 5 (52).

Figura 5. Plot de discriminación alélica basada en la emisión de fluorescencia emitida por las sondas



### **7.5 Análisis estadístico.**

Se determinó el equilibrio de Hardy-Weinberg (HWE, por sus siglas en inglés) mediante el programa FINNETI (<http://ihg2.helmholtz-muchen.de/cgi-bin/hw/hwa1>).

Utilizando el programa SPSS versión 21, Y la prueba de  $\chi^2$ , tomándose como estadísticamente significativa en valores de  $p < 0.05$ .

## 8. RESULTADOS

### 8.1 Características clínicas y demográficas de la población analizada.

En este trabajo de investigación se incluyeron 762 individuos no relacionados: 478 fueron pacientes con diagnóstico clínico de LLA, de los cuales 197 (41.2%) fueron del género femenino y 281 (58.79%) del género masculino. La población control estuvo conformada por 284 individuos sanos, 124 (43.6%) mujeres y 160 (56.3%) varones (Tabla 4). Toda la población fue <18 años de edad. Los pacientes tuvieron una edad media de  $7.9 \pm 4.7$  años y el grupo control de  $6.2 \pm 4.6$ .

Dentro de las características clínicas, el 87% de los pacientes presentaron leucemia del linaje pre-B, 7.5% de células T, 10% de carácter bifenotípico y 0.42% de células B maduras. Bajo la clasificación de FAB, el 60.25% de los individuos fueron L1, 37.45% L2 y 0.42% L3. Esta población fue estratificada por el riesgo a recaída en 4 grupos, resultando 32.01% para riesgo habitual, 1.67% riesgo intermedio, la mayor parte de la población para riesgo alto con el 63.39% y 1.05% para riesgo muy alto

Tabla 4. Características de la población analizada.

CARACTERISTICA		CASOS n (%)	CONTROLES n (%)
SEXO	F	197 (41.2)	124 (43.6)
	M	281 (58.79)	160 (56.3)
EDAD	<1	31 (6.49)	50 (17.67)
	1-10	278 (58.16)	168 (59.36)
	>10	167 (34.94)	61 (21.55)
FAB	L1	288 (60.25)	-
	L2	179 (37.45)	-
	L3	2 (0.42)	-

	SD	9 (1.88)	-
	PRE-B	419 (87.6)	-
	CEL T	38 (7.95)	-
<b>INMUNOFENOTIPO</b>	CEL B madura	2 (0.42)	-
	Bifenotípica	10 (2.09)	-
	SD	9 (1.88)	-
<b>RIESGO</b>	Habitual	153 (32.01)	-
	Intermedio	8 (1.67)	-
	Alto	303 (63.39)	-
	Muy alto	5 (1.05)	-
	SD	9 (1.88)	-

FAB: Clasificación Franco-Americano-Británico, SD: sin determinar

## 8.2 Análisis del rs4646903 en el gen *CYP1A1*.

Las distribuciones de los genotipos en la población se encuentran en HWE, mediante una prueba de chi cuadrada se encontró que hay una diferencia estadísticamente significativa en las frecuencias genotípicas entre los pacientes con respecto al grupo control ( $p=0.000723$ ) (Tabla 5).

Por otro lado, el alelo menor "C" se encontró en el 54 % de los pacientes y en aproximadamente el 41 % de los individuos control. El estudio reveló que existe asociación entre el polimorfismo, rs4646903 y la LLA (OR= 1.746, IC 95% 1.294-2.358,  $p=0.00025$ ). Además, se observó que esta variante está asociada a LLA como genotipo heterocigoto "CT" (OR= 2.118, IC 95%1.295-3.464,  $p= 0.00254$ ), bajo un modelo aditivo (OR=2.333, IC 95% 1.478-3.681:  $p= 0.00023$ ) y confiriendo mayor riesgo como homocigoto "CC" (OR=2.788, IC 95% 1.554-5.002,  $p=0.00048$ )(Tabla 5).

Tabla 5. Análisis de asociación entre el polimorfismo del gen *CYP1A1* y leucemia linfoblástica aguda en pacientes mexicanos

SNP	Genotipo/ Alelos	Casos* n (%)	Controles n (%)	OR [IC 95%],P valúe
rs4646903	TT	66(20.82)	46 (38.02)	P=0.000723
	TC	155(48.90)	51 (42.15)	
	CC	96 (30.28)	24 (10.83)	
	T	287(45.27)	143 (59.09)	1.746 [1.294-2.358] p=0.00025
	C	347 (54.73)	99 (40.91)	
	TT vs TC			2.118 [1.295-3.464] p=0.00254
	TT vs CC			2.788 [1.554-5.002] p=0.00048
TT vs TC+CC			2.333 [1.478-3.681] p=0.00023	

\* derivado de la crisis sanitaria, algunas muestras no fueron genotipificadas.

### 8.3 Análisis del rs2031920 y el rs3813867 en el gen *CYP2E1*

Del gen *CYP2E1* se analizaron dos variantes. La distribución genotípica de la variante rs2031920 no se encontró en HWE en el grupo control. Al ser comparada esta frecuencia con la distribución alélica de la base de datos de los 1000 genomas, se obtuvo que existe una diferencia significativa entre la distribución reportada y la obtenida en este estudio. (p=0.02333563).



En cuanto a la variante rs3813867, está se encontró en HWE tanto en casos como en controles, en la distribución alélica no mostraron diferencias estadísticamente significativas entre los grupos evaluados (Tabla 6).

Tabla 6. Análisis de asociación entre los polimorfismos del gen *CYP2E1* y leucemia linfoblástica aguda en pacientes mexicanos.

SNP	Genotipo/ Alelos	Casos n (%)	Controles n (%)	OR [IC 95%],P valué
<b>rs2031920</b>	CC	227 (65.04)	95 (51.35)	P= 0.007995
	CT	115 (32.95)	86 (46.49)	
	TT	7 (2.01)	4 (2.16)	
	C	569 (81.52)	276 (74.59)	0.67 [0.492-0.901] p=0.00807
	T	129 (18.48)	96 (25.41)	
	CC vs CT			0.56 [0.387-0.809] p=0.00190
	CC vs CT vs TT			0.57 [0.395-0.815] p=0.00209
<b>rs3813867</b>	GG	296 (64.49)	185 (66.79)	p=0.523121
	GC	151 (32.90)	88 (31.77)	
	CC	12 (2.61)	4 (1.44)	p=0.40545
	G	743 (52)	458 (53)	
	C	175 (10.06)	96 (17)	

#### 8.4 Análisis de los SNPs en gen *NAT2*

La distribución de los genotipos del polimorfismo rs1208 (A/G) se observó en HWE en los dos grupos de estudio, para las variaciones: rs1801280 (T/C), rs1799929 (C/T) y rs1799930 (G/A) la distribución estaba en HWE únicamente en los individuos control.

Con respecto a las variantes rs1041983 (C/T) y rs1799931 (G/A), las distribuciones de genotipos de los grupos control y de casos, no se encontraron en HWE. La comparación de las frecuencias alélicas del grupo control obtenidas con respecto a las frecuencias alélicas reportadas en la base de datos los 1000 genomas, no mostraron diferencia estadísticamente significativa para ninguna de las dos variantes (rs1041983;  $p=0.8268$  y rs1799931;  $p=0.0909$ ).

En la variante rs1041983 la distribución de los genotipos no muestra diferencias estadísticamente significativas ( $p=0.142$ ) con LLA, pero sí en la distribución alélica (OR= 0.791; IC 95% 0.629-0.996,  $p=0.04577$ ). Además, el análisis revela que la presencia de una de las copias del alelo menor confiere protección con respecto a la susceptibilidad de desarrollar LLA, mostrando bajo un modelo aditivo diferencia estadísticamente significativa (CC vs CT+TT, OR= 0.744, IC 95% 0.554-1.000  $p=0.04946$ ) (Tabla 7).

El estudio reveló que existe asociación entre LLA y el rs1801280 (OR=1.534; IC 95% 1.213-1.939,  $p=0.00034$ ) y rs1799929 (OR= 1.963; IC95% 1.549-2.487,  $p=1.74e-08$ ), En ambos casos la presencia en un solo alelo es d riesgo con respecto al desarrollo de la enfermedad (rs1801280, OR=1.443; IC95 % 1.049-1.986,  $p=0.00034$ ) y rs1799929 (C/T) (OR=1.659; IC 95% 1.206-2.283,  $p=0.00181$ ), y este riesgo aumentó en los casos homocigotos para las variantes rs1801280 (OR=2.200, IC 95% 1.304-3.710,  $p=0.00265$ ) y rs1799929 (OR=3.883, IC 95% 2.215-6.809,  $p=6.607e-07$ ) (Tabla 7).

La variante rs1799930 muestra una diferencia estadísticamente significativa en la distribución genotípica entre casos y controles ( $P=0.027980$ ), pero no se observaron

diferencias en la distribución alélica ( $p=0.370$ ). Para este SNP el genotipo GA fue más frecuente en controles que en casos (OR= 0.68; IC 95% 0.467-0.997,  $p=0.04751$ )

En cuanto a la variante rs1208 (A/G), éste muestra diferencias entre casos y controles en relación a la distribución de genotipos ( $p=0.005444$ ) y alelos (OR=1.46; IC 95% 1.16-1.83,  $p=0.0012$ ). Así mismo el genotipo homocigoto para el alelo menor GG se encontró asociado con LLA mostrando un OR= 2.32; IC 95%,  $p=0.00167$ . bajo un modelo aditivo (OR=1.45; IC 95% 1.08-1.95,  $p=0.01361$ ) (Tabla 7).

Finalmente, para la variante rs1799931 se encontraron diferencias estadísticamente significativas en las frecuencias de los genotipos entre el grupo de pacientes y de niños sanos ( $p=0.028377$ ), pero no en la distribución de los alelos ( $p=0.203$ ). Se observó que el genotipo homocigoto del alelo menor AA, protege contra LLA (OR= 0.55; IC 95%,  $p=0.02623$ ) (Tabla7).

Tabla 7. Análisis de asociación de los polimorfismos del gen *NAT2* y leucemia linfoblástica aguda en niños mexicanos.

SNP	Genotipo/ Alelos	Casos n(%)	Controles n(%)	OR [IC 95%], P valúe
<b>rs1041983</b>	CC	276 (57.9)	170 (50.5)	$p=0.142$
	CT	154 (32.3)	95 (37.5)	
	TT	47 (9.9)	17 (12.0)	
	C	706 (74.0)	435 (69.3)	0.791 [0.629-0.996] $p=0.04577$
	T	248 (26.0)	129 (30.7)	

				0.744 [0.554-1.000] p=0.04946
	<i>CC vs CT+TT</i>			
<b>rs1801280</b>	TT	225 (47.2)	165 (58.3)	p=0.003280
	TC	185 (38.8)	94 (33.2)	
	CC	66 (13.8)	22 (7.8)	
	T	635 (66.7)	424 (75.4)	1.534 [1.213-1.939] p=0.00034
	C	317 (33.3)	138 (24.6)	
	<i>TT vs TC</i>			1.443 [1.049-1.986] p=0.02408
	<i>TT vs CC</i>			2.200 [1.304-3.710] p=0.00265
	<i>TT vs TC+CC</i>			1.587 [1.178-2.138] p=0.00232
<b>rs1799929</b>	CC	206 (43.2)	170 (60.1)	p=6.9409E-7
	CT	191 (40.0)	95 (33.6)	
	TT	80 (16.8)	17 (6.0)	
	C	603 (63.2)	435 (39)	1.963 [1.549-2.487] p=1.740e-08
	T	351 (36.8)	129 (22.9)	
	<i>CC vs CT</i>			1.659 [1.206-2.283] p=0.00181
	<i>CC vs TT</i>			3.883 [2.215-6.809] p=6.607e-07
	<i>CC vs CT+TT</i>			1.997 [1.480-2.694] p=5.305e-06
<b>rs1799930</b>	GG	394 (53)	221 (54)	p=0.027980
	GA	73 (15.3)	60 (21.2)	
	AA	9 (1.9)	1 (0.4)	
	G	861 (55)	502 (89.0)	p=0.370
	A	91 (9.6)	62 (11.0)	

				0.68 [0.467-0.997] p=0.04751
<b>rs1208</b>	GG vs GA			
	AA	207 (43.4)	148 (52.3)	p=0.005444
	AG	200 (41.9)	111 (39.2)	
	GG	68 (14.3)	21 (7.4)	
	A	614 (64.6)	407 (72.7)	1.46 [1.16-1.83] p=0.0012
	G	336 (35.4)	153 (27.3)	
	AA vs GG			2.32 [1.36-3.95] p=0.00167
	AA vs AG+ GG			1.45 [1.08-1.95] p=0.01361
<b>rs1799931</b>	GG vs AA			p=0.028377
	GG	328 (56)	192 (57)	
	GA	117 (24.5)	56 (19.8)	
	AA	29 (6.1)	31 (11)	
	G	773 (81.5)	440 (54)	p=0.203
	A	175 (18.5)	118 (21.1)	
	GG vs AA			0.55 [0.32-0.94] p=0.02623

OR: Odds Ratio, IC: intervalo de confianza

### 8.5 Análisis del rs1800566 del gen *NQO1*

En el análisis del rs1800566, las frecuencias genotípicas se encontraban en HWE. En cuanto al análisis de asociación entre casos y controles, no se observaron diferencias estadísticamente significativas al comparar las frecuencias genotípicas (p=0.423) y alélicas (p=0.291) (Tabla 8).

Tabla 8. Análisis de asociación entre el polimorfismo del gen *NQO1* y leucemia linfoblástica aguda en niños mexicanos.

SNP	Genotipo / Alelos	Casos n (%)	Controles n (%)	P
Rs1800566	CC	156 (32.8)	79 (28.2)	p= 0.423
	CT	228 (47.9)	144 (51.4)	
	TT	92 (19.3)	57 (20.4)	
Rs1800566	C	540 (56.7)	302(53.9)	p=0.291
	T	412 (43.3)	258 (46.1)	

## 9. DISCUSIÓN

La leucemia linfoblástica aguda (LLA) es la neoplasia hematológica más frecuente en edad pediátrica a nivel mundial, actualmente la supervivencia global alcanza hasta el 90% en países desarrollados, sin embargo, difiere del escenario visto en países en vías de desarrollo como México, donde las cifras son menores al 60% (53, 54). La LLA es una entidad compleja cuya etiología no se encuentra del todo clara, aunque se sabe que esta enfermedad surge de la interacción de diversos factores, tanto genéticos como ambientales. Dentro de estos últimos se encuentra la exposición a sustancias exógenas como: pesticidas, humo de cigarro, pintura, medicamentos, compuestos aromáticos, etc (16, 55).

En el entendimiento de la relación gen-ambiente y LLA, una de las vías que más se ha estudiado es el metabolismo de xenobioticos y los polimorfismos en los genes

que codifican a las enzimas involucradas en su biotransformación, Se ha reportado que algunos polimorfismos pueden explicar las diferencias interindividuales frente a sustancias tóxicas y en su capacidad de metabolizar o activar sustancias carcinogénicas (56, 57). Dentro del contexto antes mencionado, en este trabajo se realizó un estudio de casos y controles con el fin de identificar la asociación entre LLA y polimorfismos de los genes *CYP1A1*, *CYP2E1*, *NQO1* y *NAT2*.

### **9.1 CYP1A1**

El gen *CYP1A1* codifica a una enzima clave en la bioactivación de xenobióticos, catalizando el primer paso del metabolismo de hidrocarburos aromáticos policíclicos, en donde se crean productos intermedios reactivos que usualmente se conjugan en la fase II, estos también pueden interactuar con el material genético uniéndose al ADN (58). La variante rs4646903 es una mutación que no altera la proteína, pero que ha sido asociada con el incremento de actividad enzimática (64). A lo largo de los años se ha asociado con el riesgo de desarrollar diferentes tipos de cáncer entre ellos la leucemia aguda, sin embargo, los datos de asociación entre esta variante con LLA, han sido controversiales ya que algunos de estos no han encontrado asociación clara entre el rs4644603 y esta enfermedad (35, 38).

Nuestros resultados muestran una asociación entre el SNP rs4646903 y LLA (OR= 1.746, IC 95% 1.294-2.358, p=0.00025), como lo reportado por Gallegos- Arreola en una población adulta mexicana, donde también se identificó al alelo "C" como una variante de riesgo para desarrollar esta enfermedad (59), un hallazgo similar fue reportado por Joseph y colaboradores, cuyo estudio se realizó en población del sur de la India(60).

Existen otros trabajos que no han encontrado asociación del rs4646903 con la LLA por si sola, pero si en conjunto con variantes de otros genes como el *MLHI* (gen que codifica a una proteína encargada de la reparación del ADN) y *GSTM* ( gen que codifica a una proteína involucrada en pasos subsecuentes a p450 en el metabolismo de xbs) (61-63). También se debe considerar el trabajo de Lee y colaboradores en donde al momento de realizar una estratificación de la población por su origen étnico, el modelo homocigoto para la variante rs4646903 se asoció con LLA en población hispana y no en el resto de la población, este último indica que puede deberse a que la población hispana que ellos estudian tienen diferentes ocupaciones que los hacen más propensos a la exposición de compuestos tóxicos y por ello la frecuencia de esta variante sea mayor (64).

Considerando los resultados que se han reportado con respecto a esta variante y la LLA, se sugiere completar la genotipificación de las muestras disponibles para este estudio con el fin de confirmar el mantenimiento de la asociación que se muestra hasta ahora o bien aumentar la población de estudio para tener una mejor descripción de esta variante en la población infantil mexicana.

## **9.2 CYP2E1**

Dentro de las isoformas del complejo p450, se encuentra el gen *CYP2E1* cuya proteína ayuda en la transformación de compuestos de bajo peso molecular como el alcohol, acetona, anestésicos, etc. Al igual que *CYP1A1*, muchos de los compuestos catalizados por *CYP2E1* forman metabolitos que se unen a DNA o son especies reactivas de oxígeno.(65) En el presente trabajo se hizo el análisis de dos



polimorfismos, el rs203190 y rs3813867, variantes que se han asociado al incremento de la actividad enzimática.(44)

Del mismo modo que la variante CYP1A1\*2A la Genotipificación del polimorfismo rs2031920 en el grupo control no se pudo completar lo cual podría explicar que la distribución genotípica no se encuentra HWE, se encontró que las frecuencias alélicas observadas en nuestro estudio son diferentes a lo reportado en el proyecto de los 1000 genomas, sin embargo, dicha diferencia no fue estadísticamente significativa ( $p=0.0233$ ), razón por la cual no se debe descartar totalmente los resultados presentes en esta tesis.

Previamente se ha reportado asociación de los SNPs de *CYP2E1* con LLA, entre ellos el trabajo de Aydin y colaboradores donde indican que hay asociación de riesgo del rs2031920 en con respecto a LLA tanto en pacientes pediátricos como adultos (65).Otros reportes indican que no hay una asociación de la variante por si sola, si no en conjunto ya sea con otras variantes en el mismo gen u otros genes como: *GSTM*, *GSTP1*, *CYP1A1*, *XRCC1*, *MLH1* (62, 63, 66) también se ha reportado asociación de *CYP2E1* con la LLA, cuando las variables se conjuntan con condiciones ambientales específicas como el consumo de alcohol o cigarro por parte de los padres (17). Considerando lo poco claro que queda el papel de este gen con la LLA, se sugiere no descartar futuros análisis de estos polimorfismos con respecto a LLA y complementar los reportado en esta tesis.

El polimorfismo rs3813867, mostró una distribución genotípica en HWE, la distribución alélica obtenida no reveló asociación entre esta variante con el riesgo

de desarrollar leucemia linfoblástica aguda, resultado que encaja con lo encontrado por Canalle y colaboradores y Bonaventure (62, 67).

### **9.3 Gen *NQO1***

*NQO1* se encuentra involucrado en el metabolismo de diferentes sustancias potencialmente cancerígenas, y se ha propuesto como una enzima anti-cáncer por su papel en la protección del daño oxidativo resultado de reacciones de reducción que producen semiquinonas, razón por la cual se han asociado a este gen con diversos tipos de cáncer como el de colón (68).

En nuestro estudio no se encontró asociación del polimorfismo rs1800566 con la LLA, este hallazgo ya ha sido reportado previamente por algunos investigadores, Sirma y colaboradores expresan no haber encontrado asociación y tampoco una diferencia significativa en las frecuencias de los alelos con respecto a otras poblaciones, algunos de los estudios también refieren a que no hay asociación de esta variante por si sola ni estratificando a la población estudiada por hábitos o respecto a otras variantes (69-72).

Nuestros resultados contrastan con otros reportes en donde si han encontrado asociación de esta variante con la LLA. Silviera y colaboradores para una estudio realizado en Brazil, encuentran que el rs1800566 se asocia como variante de protección con respecto a la LLA, mismo fenómeno fue observado por Yeoh en población asiática, por otro lado estudios realizados en filipinos y en Canadienses revelan que esta variante confiere riesgo a desarrollar LLA por si sola y en conjunto con *GSTM* (73-75). Considerando dichas discrepancias, podría atribuirse el

comportamiento de esta variante con respecto al origen étnico o incluso a otros fenómenos que puedan estar afectando la frecuencia de los alelos, en lo que respecta a este trabajo, se ha descrito al alelo rs1800566 en una población infantil mexicana con respecto a la leucemia linfoblástica aguda.

#### **9.4 Gen *NAT2***

El gen *NAT2* codifica a una proteína de fase II del metabolismo de Xbs, cuya actividad modifica la predisposición en el desarrollo del cáncer por su rol en la bioactivación y detoxificación de carcinogénicos aromáticos o amino heterocíclicos, su actividad radica en la transferencia de un grupo acetilo a dichos compuestos o sus metabolitos (46). *NAT2* es un gen altamente polimórfico, cuyas variantes se han asociado previamente con diversos tipos de cáncer como: cáncer de mama, pulmón, vejiga y leucemia aguda (76). En la presente tesis se estudiaron algunas de las variantes de este gen, cuatro en donde la mutación produce un cambio de aminoácido: rs1801280 (T341C, Ile114Thr), rs1799930 (G590A, Arg197Gln), rs1799931 (G857A, Lys286Glu), rs1208 (A803G, Lys268Arg) y dos cuya mutación no altera a la proteína: rs1799929 (C481T) y rs1041983 (C282T).

En lo que refiere a los polimorfismos de *NAT2*, las frecuencias obtenidas para el alelo mutante en niños mexicanos fueron de 0.307 para el rs1041983, 0.246 para rs1801280, 0.229 del rs1799929, 0.11 del rs1799930, 0.273 para rs1208 y 0.211 para rs1799931. Las frecuencias alélicas encontradas de todas las variantes difieren de aquellas que se reportaron con anterioridad para población mexicana por Taja- Chayeb y colaboradores, quienes evaluaron adultos de población abierta (77), considerando como una posibilidad que la estratificación poblacional sea un factor

que explique esta diferencia. En este sentido, algunos estudios han evidenciado que el entorno puede tener un impacto importante en las frecuencias alélicas para algunos genes ya que la exposición a ciertos xenobióticos puede brindar un nuevo régimen de selectividad favoreciendo el mantenimiento de algunas variantes sobre otras (78).

En *NAT2*, se encontró asociación estadísticamente significativa como variantes de riesgo para el desarrollo de LLA, a los polimorfismos rs1801280, rs1799929 y al rs1208, en donde el riesgo es mayor bajo un modelo recesivo. (OR=2.20; IC95% 1.304-3.7110, OR=3.883; IC95%2.21-6.809 y OR=2.32; IC95%1.36-3.95, respectivamente).

En estudios anteriores se ha reportado que hay asociación de riesgo para la variable rs1801280 en una población infantil brasileña. Este mismo hallazgo fue reportado más adelante en el meta-análisis realizado por Zhu y colaboradores, en donde además encontraron que esta variante mantiene su asociación al momento de estratificar por su origen étnico (76, 79). Por el contrario, Kamel y colaboradores no encuentran asociación para esta variable por sí sola, pero sí al hacer análisis con respecto a la clasificación fenotípica de acetiladores, en donde T341C se encuentra involucrado (18). En el caso del rs1799929, nuestros datos corresponden con lo reportado en los estudios antes mencionados, en donde se asoció a este SNP como variable de riesgo.

Finalmente, en el caso de la variable rs1208, no se ha reportado una asociación con respecto a la LLA de manera individual, pero si en conjunto con otras variantes

tomando a los fenotipos de acetiladores, principalmente aquellos que involucran a las variantes C431T y C481T (76).

Dentro de las variantes estudiadas para el gen *NAT2*, el rs1041983 (C282T) mostró al alelo menor "T" como variante de protección, la asociación para esta variante también se encontró bajo un modelo aditivo con un valor de OR=0.774; IC95% 0.554-1.00,  $p=0.04946$ , resultado que difiere totalmente de lo encontrado por Hernández y colaboradores en un estudio realizado en infantes mexicanos originarios de San Luis potosí, donde ubican a esta como una variable de riesgo(80). Algunos otros estudios no han encontrado una asociación de la variable sola, pero sí, si se encuentra involucrada con las demás variables en análisis por fenotipos de acetiladores principalmente en conjunto con la variable rs1799930, cuya asociación es de riesgo(6). Esta discrepancia puede explicarse por la diferencia de la población tanto en tamaño como región, por lo cual es aconsejable revisar la reproducibilidad en las frecuencias de estas variantes o bien hacer una ampliación del estudio que contemple diferentes regiones para mejorar la descripción de SNP en población pediátrica de México.

El rs1799930 se encontró que está asociada a la LLA asociada bajo un modelo dominante (OR=0.68; IC95% 0.467-0.997,  $p=0.04751$ ). En el estudio realizado por Hernández se reportó una asociación como variable de riesgo, al igual que lo encontrado por Kamel y colaboradores(80) (18) Como se mencionó previamente la frecuencia de ciertas variantes puede verse afectada por otros factores, por la cual se sugiere validar estos hallazgos en poblaciones independientes.

G857A o rs1799931 se asoció como variable de protección bajo un modelo recesivo (OR=0.55; IC95% 0.32-0.94,  $p=0.02623$ ), este mismo comportamiento fue observado previamente para una población mestiza y para asiáticos, pero diferente al estudio en población pediátrica infantil mexicana(6) (80).

Aunado a lo anterior, estudios han reportado que la actividad de NAT2 puede ser afectada por otros factores como: dieta, medicación tradicional, entorno laboral, además del origen étnico (78, 81), tomando esto en cuenta es posible que nuestros resultados manifiesten la influencia de estos factores, razón por la cual no coincidan del todo con los datos previamente reportados.

Dentro de las limitaciones de este trabajo fue la falta de tiempo para realizar una parte del proceso experimental causado por el cierre temporal del laboratorio, además de la cantidad relativamente menor de muestras control, un aumento en la N de este grupo permitirá corroborar el comportamiento de las frecuencias alélicas para algunos SNPs, además de esclarecer los datos de asociación para aquellos SNPs cuyos intervalos de confianza o p valúe se encuentran por el límite de significancia (82).

En resumen, este trabajo permitió conocer las frecuencias de los polimorfismos de los genes de *CYP2E1*, *NQO1* y *NAT2*, en niños mexicanos, e identificar que SPNs en el gen *NAT2* se asocian a LLA. A pesar de no haber encontrado asociación entre LLA y los polimorfismos rs3813867, rs1800566, no se puede descartar la contribución de los genes *CYP2E1* y *NQO1* al desarrollo de la leucemia linfoblástica aguda, puesto que las variantes estudiadas no son únicas en estos, además diversos estudios han reportado que algunos polimorfismos pueden contribuir como

variantes de riesgo o protección en conjunto con alteraciones en otros genes. Nuestros resultados fueron publicados en *Frontiers Oncology* (anexo I) (83).

## 10. CONCLUSIONES

El desarrollo de este proyecto ha permitido identificar puntos importantes con respecto a la leucemia linfoblástica aguda y polimorfismos en genes involucrados en la biotransformación de xenobióticos en población infantil mexicana, por lo cual se puede concluir lo siguiente:

1. Existen diferencias en la distribución de alelos en los genes *CYP1A1*, *CYP2E1* y *NAT2*.
2. Los polimorfismos rs3813867 del *CYP2E1* y rs1800566 del gen *NQO1*, no se encuentran asociadas con el riesgo de desarrollar leucemia linfoblástica aguda en niños mexicanos.
3. Polimorfismos en el gen *NAT2* se asocian con la susceptibilidad a desarrollar LLA en niños mexicanos

## 11. PERSPECTIVAS

- Concluir las genotipificaciones para las variantes de los genes *CYP1A1* y *CYP2E1*
- Validar los hallazgos obtenidos en el presente trabajo.
- Realizar análisis de asociación usando en conjuntos las variantes de los genes *CYP1A1*, *CYP2E1* Y *NAT2*.
- Aumentar la N de estudio, considerando más regiones de México.

## 12. REFERENCIAS

1. Shalkow J NS, Palomares G, Esparza M, Castro L, JAimes A, et al. Comportamiento Epidemiológico del Cáncer en menores de 18 años. México 2008-2014. Secretaria de Salud 2016. p. 29-38.
2. Lam P, Steinmetz NF. Delivery of siRNA therapeutics using cowpea chlorotic mottle virus-like particles. *Biomaterials science*. 2019;7(8):3138-42.
3. Gobierno de México: Centro Nacional para la Salud de la Infancia y Adolescencia [Internet]. México: Gobierno de México 2019. [cited 2019]. Available from: <https://www.gob.mx/salud/censia/articulos/cancer-infantil-en-mexico-130956?state=published>.
4. Lassaletta AA. Leucemias. Leucemia linfoblástica aguda *Pediatría integral* 2016;6(XX):380-9.
5. Jiménez S, Hidalgo A, J R. Leucemia linfoblástica aguda infantil: una aproximación genómica *Boletín Médico del Hospital Infantil de México*. 2017;74(1):13-26.
6. Zanrosso CW, Emerenciano M, Faro A, Gonçalves BA, Mansur MB, Pombo-de-Oliveira MS. Genetic variability in N-acetyltransferase 2 gene determines susceptibility to childhood lymphoid or myeloid leukemia in Brazil. *Leuk Lymphoma*. 2012;53(2):323-7.
7. Schoonen L, Eising S, van Eldijk MB, Bresseleers J, Van Der Pijl M, Nolte RJ, et al. Modular, bioorthogonal strategy for the controlled loading of cargo into a protein nanocage. *Bioconjugate chemistry*. 2018;29(4):1186-93.
8. Williams L. Acute Lymphoblastic Leukemias. *Clinical Laboratory Hematology*. 2nd ed. San Antonio: Pearson; 2010. p. 692-709.
9. Chiaretti S, Zini G, Bassan R. Diagnosis and subclassification of acute lymphoblastic leukemia. *Mediterr J Hematol Infect Dis*. 2014;6(1):e2014073.
10. Sala ML, Blanco B, Pérez M. 10. Hematología clínica. *Sociedad Española de Farmacia Hospitalaria*. p. 1031-76.
11. Carroll W, Bhatla T. Acute lymphoblastic leukemia. *Lanzkowsky's Manual of Pediatric Hematology and Oncology*: Elsevier; 2016. p. 367-89.



12. Rurup WF, Verbij F, Koay MS, Blum C, Subramaniam V, Cornelissen JJ. Predicting the loading of virus-like particles with fluorescent proteins. *Biomacromolecules*. 2014;15(2):558-63.
13. Arriaga-Pizano L, Ramírez-Ramírez D, Prieto-Chávez J, Pelayo R, Ruiz-Argüelles A. Reporte de la Primera Reunión Nacional de Consenso para la Inmunofenotipificación de Leucemias Agudas. *Gaceta medica de Mexico*. 2019;155(1):20-9.
14. Michel G. Leucemia linfoblástica aguda del niño y del adolescente: clínica y tratamiento. *EMC-Pediatría*. 2008;43(4):1-11.
15. Dorantes E, Medina A, Dávila K, López B. Clasificación inmunológica de las leucemias agudas linfoblásticas del Hospital Infantil de México Federico Gómez, de acuerdo al EGIL (European Group for the Immunological Classification of Leukemia). *Gac Mexicana Oncol*. 2013;12(3):136-42.
16. Hitzler J, Jiménez E, Jongmans M, López B, al. e. Etiology of Acute Leukemias in Children. Ciudad de México Springer 2016. Available from: <https://www.springer.com/gp/book/9783319057972>.
17. Infante Rivard C, Krajcinovic M, Labuda D, Sinnett D. Childhood acute lymphoblastic leukemia associated with parental alcohol consumption and polymorphisms of carcinogen-metabolizing genes. *Epidemiology*. 2002:277-81.
18. Kamel AM, Ebid GT, Moussa HS. N-Acetyltransferase 2 (NAT2) polymorphism as a risk modifier of susceptibility to pediatric acute lymphoblastic leukemia. *Tumor Biology*. 2015;36(8):6341-8.
19. Yaqoob SB, Adnan R, Rameez Khan RM, Rashid M. Gold, Silver, and Palladium Nanoparticles: A Chemical Tool for Biomedical Applications. *Frontiers in chemistry*. 2020;8:376.
20. Mullighan CG. Genomic characterization of childhood acute lymphoblastic leukemia. *Semin Hematol*. 2013;50(4):314-24.
21. Roberts KG, Mullighan CG. Genomics in acute lymphoblastic leukaemia: insights and treatment implications. *Nat Rev Clin Oncol*. 2015;12(6):344-57.
22. McInnes RR, Willard HF, Nussbaum R. Thompson & Thompson *Genética Médica*: Elsevier Brasil; 2016.

23. Caratachea MAC. Polimorfismos genéticos: Importancia y aplicaciones. *Revista del Instituto Nacional de Enfermedades Respiratorias*. 2007;20(3):213-21.
24. Newton-Cheh C, Hirschhorn JN. Genetic association studies of complex traits: design and analysis issues. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. 2005;573(1-2):54-69.
25. Foulkes AS. *Applied statistical genetics with R*: Springer; 2009.
26. Panoutsopoulou K, Wheeler E. Key concepts in genetic epidemiology. *Genetic Epidemiology*: Springer; 2018. p. 7-24.
27. Song JW, Chung KC. Observational studies: cohort and case-control studies. *Plastic and reconstructive surgery*. 2010;126(6):2234.
28. Laird NM, Lange C. Family-based designs in the age of large-scale gene-association studies. *Nature Reviews Genetics*. 2006;7(5):385-94.
29. Manikandan P, Nagini S. Cytochrome P450 structure, function and clinical significance: a review. *Current drug targets*. 2018;19(1):38-54.
30. González JCR, Guerra IR. El sistema citocromo P450 y el metabolismo de xenobióticos. *Revista Cubana de Farmacia*. 2014;48(3):495-507.
31. Almazroo OA, Miah MK, Venkataramanan R. Drug metabolism in the liver. *Clinics in liver disease*. 2017;21(1):1-20.
32. Anzenbacher P, Zanger UM. *Metabolism of Drugs and Other Xenobiotics*: Wiley; 2012.
33. Information NCfB. rs4646903 [Homo sapiens]: U.S. National Library of Medicine: dbSNP short genetic variations; 2019 [Available from: <https://www.ncbi.nlm.nih.gov/snp/rs4646903#publications>].
34. Ouerhani S, Cherif N, Bahri I, Safra I, Menif S, Abbes S. Genetic polymorphisms of NQO1, CYP1A1 and TPMT and susceptibility to acute lymphoblastic leukemia in a Tunisian population. *Molecular biology reports*. 2013;40(2):1307-14.
35. Zhuo W, Zhang L, Qiu Z, Zhu B, Chen Z. Does cytochrome P450 1A1 MspI polymorphism increase acute lymphoblastic leukemia risk? Evidence from 2013 cases and 2903 controls. *Gene*. 2012;510(1):14-21.

36. Quiñones L, Lee K, Escala M, García K, Godoy L, Castro A, et al. Farmacogenética del cáncer: Estudio de variaciones genéticamente determinadas en la susceptibilidad a cáncer por exposición a xenobióticos. *Rev Med Chile*. 2006;134:499-515.
37. Swinney RM, Beuten J, Collier AB, Chen TT-L, Winick NJ, Pollock BH, et al. Polymorphisms in CYP1A1 and ethnic-specific susceptibility to acute lymphoblastic leukemia in children. *Cancer Epidemiology and Prevention Biomarkers*. 2011;20(7):1537-42.
38. Lu J, Zhao Q, Zhai Y-J, He H-R, Yang L-H, Gao F, et al. Genetic polymorphisms of CYP1A1 and risk of leukemia: a meta-analysis. *OncoTargets and therapy*. 2015;8:2883.
39. Nida S, Javid B, Akbar M, Idrees S, Adil W, Ahmad GB. Gene variants of CYP1A1 and CYP2D6 and the risk of childhood acute lymphoblastic leukaemia; outcome of a case control study from Kashmir, India. *Molecular Biology Research Communications*. 2017;6(2):77.
40. Chen J, Jiang S, Wang J, Renukuntla J, Sirimulla S, Chen J. A comprehensive review of cytochrome P450 2E1 for xenobiotic metabolism. *Drug metabolism reviews*. 2019;51(2):178-95.
41. Infante Rivard C, Amre D, Sinnett D. GSTT1 and CYP2E1 polymorphisms and trihalomethanes in drinking water: effect on childhood leukemia. *Environ Health Perspect*. 2002;110(6):591-3.
42. Gene USA: National Center for Biotechnology information; 2019 [Available from: <https://www.ncbi.nlm.nih.gov/gene/>].
43. Krajcinovic M, Sinnett H, Richer C, Labuda D, Sinnett D. Role of NQO1, MPO and CYP2E1 genetic polymorphisms in the susceptibility to childhood acute lymphoblastic leukemia. *International journal of cancer*. 2002;97(2):230-6.
44. Ulusoy G, Adali O, Tumer TB, Sahin G, Gozdasoglu S, Arinç E. Significance of genetic polymorphisms at multiple loci of CYP2E1 in the risk of development of childhood acute lymphoblastic leukemia. *Oncology*. 2007;72(1-2):125-31.
45. Rodríguez EMdR. Defensa química y citocromo P450: relación con la defensa inmune. *Revista Médica de la Universidad Veracruzana*. 2011;11(2):53-63.

46. Walraven JM, Zang Y, Trent JO, Hein DW. Structure/function evaluations of single nucleotide polymorphisms in human N-acetyltransferase 2. *Current drug metabolism*. 2008;9(6):471-86.
47. Nebert DW, Roe AL, Vandale SE, Bingham E, Oakley GG. NAD (P) H: quinone oxidoreductase (NQO1) polymorphism, exposure to benzene, and predisposition to disease: a HuGE review. *Genetics in Medicine*. 2002;4(2):62-70.
48. Zaker F, Safaei A, Nasiri N, Abdollahzadeh M, Pazhakh V. The association of NAD (P) H: quinone oxidoreductase gene polymorphisms with pediatric acute lymphoblastic leukemia. *Laboratory Medicine*. 2012;43(6):256-61.
49. Park J-S, Lee Y-Y, Kim J, Seo H, Kim H-S.  $\beta$ -Lapachone increases phase II antioxidant enzyme expression via NQO1-AMPK/PI3K-Nrf2/ARE signaling in rat primary astrocytes. *Free Radical Biology and Medicine*. 2016;97:168-78.
50. He H, Zhai X, Liu X, Zheng J, Zhai Y, Gao F, et al. Associations of NQO1 C609T and NQO1 C465T polymorphisms with acute leukemia risk: a PRISMA-compliant meta-analysis. *OncoTargets and therapy*. 2017;10:1793.
51. Guha N, Chang JS, Chokkalingam AP, Wiemels JL, Smith MT, Buffler PA. NQO1 polymorphisms and de novo childhood leukemia: a HuGE review and meta-analysis. *American journal of epidemiology*. 2008;168(11):1221-32.
52. Mireya RF. Análisis de alelos deficientes del gen tiopurina s-metiltransferasa (tpmt) en pacientes pediátricos con leucemia aguda linfoblástica: Universidad Nacional Autónoma de México; 2016.
53. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, 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. *The Lancet*. 2018;391(10125):1023-75.
54. Bárcenas D. Identificación de genes diferencialmente expresados entre pacientes con leucemia aguda linfoblástica de bajo y de alto riesgo. México: Universidad Nacional Autónoma de México; 2018.
55. Chokkalingam AP, Buffler PA. Genetic susceptibility to childhood leukaemia. *Radiation protection dosimetry*. 2008;132(2):119-29.

56. Nousome D, Lupo PJ, Okcu MF, Scheurer ME. Maternal and offspring xenobiotic metabolism haplotypes and the risk of childhood acute lymphoblastic leukemia. *Leukemia research*. 2013;37(5):531-5.
57. Kreile M, Piekuse L, Rots D, Dobele Z, Kovalova Z, Lace B. Analysis of possible genetic risk factors contributing to development of childhood acute lymphoblastic leukaemia in the Latvian population. *Archives of medical science: AMS*. 2016;12(3):479.
58. Masson LF, Sharp L, Cotton SC, Little J. Cytochrome P-450 1A1 gene polymorphisms and risk of breast cancer: a HuGE review. *American journal of epidemiology*. 2005;161(10):901-15.
59. Gallegos-Arreola M, González-García J, Figuera L, Puebla-Pérez A, Delgado-Lamas J, Zúñiga-González G. Distribution of CYP1A1\* 2A polymorphism in adult patients with acute lymphoblastic leukemia in a Mexican population. *Blood Cells, Molecules, and Diseases*. 2008;41(1):91-4.
60. Joseph T, Kusumakumary P, Chacko P, Abraham A, Radhakrishna Pillai M. Genetic polymorphism of CYP1A1, CYP2D6, GSTM1 and GSTT1 and susceptibility to acute lymphoblastic leukaemia in Indian children. *Pediatric blood & cancer*. 2004;43(5):560-7.
61. Krajinovic M, Labuda D, Richer C, Karimi S, Sinnett D. Susceptibility to childhood acute lymphoblastic leukemia: influence of CYP1A1, CYP2D6, GSTM1, and GSTT1 genetic polymorphisms. *Blood*. 1999;93(5):1496-501.
62. Canalle R, Burim RV, Tone LG, Takahashi CS. Genetic polymorphisms and susceptibility to childhood acute lymphoblastic leukemia. *Environmental and molecular mutagenesis*. 2004;43(2):100-9.
63. Mathonnet G, Krajinovic M, Labuda D, Sinnett D. Role of DNA mismatch repair genetic polymorphisms in the risk of childhood acute lymphoblastic leukaemia. *British Journal of Haematology*. 2003;123(1):45-8.
64. Lee K-M, Ward MH, Han S, Ahn HS, Kang HJ, Choi HS, et al. Paternal smoking, genetic polymorphisms in CYP1A1 and childhood leukemia risk. *Leukemia Research*. 2009;33(2):250-8.

65. Aydin-Sayitoglu M, Hatirnaz O, Erensoy N, Ozbek U. Role of CYP2D6, CYP1A1, CYP2E1, GSTT1, and GSTM1 genes in the susceptibility to acute leukemias. *American journal of hematology*. 2006;81(3):162-70.
66. Tumer TB, Yilmaz D, Tanrikut C, Sahin G, Ulusoy G, Arinç E. DNA repair XRCC1 Arg399Gln polymorphism alone, and in combination with CYP2E1 polymorphisms significantly contribute to the risk of development of childhood acute lymphoblastic leukemia. *Leuk Res*. 2010;34(10):1275-81.
67. Bonaventure A, Goujon-Bellec S, Rudant J, Orsi L, Leverger G, Baruchel A, et al. Maternal smoking during pregnancy, genetic polymorphisms of metabolic enzymes, and childhood acute leukemia: the ESCALE study (SFCE). *Cancer Causes & Control*. 2012;23(2):329-45.
68. Ding R, Lin S, Chen D. Association of NQO1 rs1800566 polymorphism and the risk of colorectal cancer: a meta-analysis. *International journal of colorectal disease*. 2012;27(7):885-92.
69. Sirma S, Agaoglu L, Yildiz I, Cayli D, Horgusluoglu E, Anak S, et al. NAD(P)H:quinone oxidoreductase 1 null genotype is not associated with pediatric de novo acute leukemia. *Pediatr Blood Cancer*. 2004;43(5):568-70.
70. Smith MT, Wang Y, Skibola CF, Slater DJ, Nigro LL, Nowell PC, et al. Low NAD(P)H:quinone oxidoreductase activity is associated with increased risk of leukemia with MLL translocations in infants and children. *Blood*. 2002;100(13):4590-3.
71. Lanciotti M, Dufour C, Corral L, Di Michele P, Pigullo S, De Rossi G, et al. Genetic polymorphism of NAD(P)H:quinone oxidoreductase is associated with an increased risk of infant acute lymphoblastic leukemia without MLL gene rearrangements. *Leukemia*. 2005;19(2):214-6.
72. De Aguiar Gonçalves BA, Vasconcelos GM, Thuler LCS, Andrade C, Faro A, Pombo-De-Oliveira MS. NQO1 rs1800566 (C609T), PON1 rs662 (Q192R), and PON1 rs854560 (L55M) polymorphisms segregate the risk of childhood acute leukemias according to age range distribution. *Cancer Causes & Control*. 2012;23(11):1811-9.

73. Rimando MG, Chua MN, Yuson E, de Castro-Bernas G, Okamoto T. Prevalence of GSTT1, GSTM1 and NQO1 (609C>T) in Filipino children with ALL (acute lymphoblastic leukaemia). *Biosci Rep.* 2008;28(3):117-24.
74. Silveira Vda S, Canalle R, Scrideli CA, Queiroz RG, Tone LG. Role of the CYP2D6, EPHX1, MPO, and NQO1 genes in the susceptibility to acute lymphoblastic leukemia in Brazilian children. *Environ Mol Mutagen.* 2010;51(1):48-56.
75. Yeoh AE-J, Lu Y, Chan JY-S, Chan YH, Ariffin H, Kham SK-Y, et al.
76. Zhu X, Liu Y, Chen G, Guo Q, Zhang Z, Zhao L, et al. Association between NAT2 polymorphisms and acute leukemia risk: A meta-analysis. *Medicine.* 2019;98(12).
77. Taja-Chayeb L, Agúndez J, Miguez-Muñoz C, Chavez-Blanco A, Dueñas-Gonzalez A. Arylamine N-acetyltransferase 2 genotypes in a Mexican population. *Genet Mol Res.* 2012;11(2):1082-92.
78. Sabbagh A, Darlu P, Crouau-Roy B, Poloni ES. Arylamine N-acetyltransferase 2 (NAT2) genetic diversity and traditional subsistence: a worldwide population survey. *PloS one.* 2011;6(4):e18507.
79. Zanrosso CW, Emerenciano M, de Aguiar Gonçalves BA, Faro A, Koifman S, Pombo-de-Oliveira MS. N-acetyltransferase 2 polymorphisms and susceptibility to infant leukemia with maternal exposure to dipyrrone during pregnancy. *Cancer Epidemiology and Prevention Biomarkers.* 2010;19(12):3037-43.
80. Hernández-González O, Ortiz-Zamudio JJ, Rodríguez-Pinal CJ, Alvarado-Morales I, Martínez-Jiménez VDC, Salazar-González RA, et al. Genetic polymorphisms of arylamine N-acetyltransferases 1 and 2 and the likelihood of developing pediatric acute lymphoblastic leukemia. *Leukemia & Lymphoma.* 2018;59(8):1968-75.
81. Djordjevic N, Carrillo JA, Roh HK, Karlsson S, Ueda N, Bertilsson L, et al. Comparison of N-acetyltransferase-2 enzyme genotype-phenotype and xanthine oxidase enzyme activity between Swedes and Koreans. *The Journal of Clinical Pharmacology.* 2012;52(10):1527-34.

82. Cerda J, Vera C, Rada G. Odds ratio: aspectos teóricos y prácticos. *Revista médica de Chile*. 2013;141(10):1329-35.
83. Medina-Sanson A, Núñez-Enríquez JC, Hurtado-Cordova E, Pérez-Saldivar ML, Martínez-García A, Jiménez-Hernández E, et al. Genotype-Environment Interaction Analysis of NQO1, CYP2E1, and NAT2 Polymorphisms and the Risk of Childhood Acute Lymphoblastic Leukemia: A Report From the Mexican Interinstitutional Group for the Identification of the Causes of Childhood Leukemia. *Front Oncol*. 2020;10:571869.





# Genotype-Environment Interaction Analysis of NQO1, CYP2E1, and NAT2 Polymorphisms and the Risk of Childhood Acute Lymphoblastic Leukemia: A Report From the Mexican Interinstitutional Group for the Identification of the Causes of Childhood Leukemia

Aurora Medina-Sanson<sup>1,2</sup>, Juan Carlos Núñez-Enriquez<sup>2</sup>, Eduardo Hurtado-Cordova<sup>4,5</sup>, María Luisa Pérez-Saldivar<sup>2</sup>, Anayeli Martínez-García<sup>4,6</sup>, Elva Jiménez-Hernández<sup>7</sup>, Juan Carlos Fernández-López<sup>8</sup>, Jorge Alfonso Martín-Trejo<sup>9</sup>, Héctor Pérez-Lorenzana<sup>10</sup>, Janet Flores-Lujano<sup>2</sup>, Raquel Amador-Sánchez<sup>11</sup>, Félix Gustavo Mora-Ríos<sup>12</sup>, José Gabriel Peñaloza-González<sup>13</sup>, David Aldebarán Duarte-Rodríguez<sup>14</sup>, José Refugio Torres-Nava<sup>15</sup>, Juan Eduardo Flores-Bautista<sup>16</sup>, Rosa Martha Espinosa-Elizondo<sup>16</sup>, Pedro Francisco Román-Zapeda<sup>17</sup>, Luz Victoria Flores-Villegas<sup>18</sup>, Juana Esther González-Ulvarrí<sup>19</sup>, Sofía Inere Martínez-Silva<sup>20</sup>, Gilberto Espinoza-Arribas<sup>21</sup>, Carolina Almeida-Hernández<sup>22</sup>, Rosario Ramírez-Colonado<sup>23</sup>, Luis Hernández-Mora<sup>24</sup>, Luis Ramiro García-López<sup>25</sup>, Gabriela Adriana Cruz-Ojeda<sup>26</sup>, Arturo Emilio Godoy-Esquivel<sup>27</sup>, Iris Contreras-Hernández<sup>28</sup>, Abraham Medina-Hernández<sup>29</sup>, María Guadalupe López-Caballero<sup>30</sup>, Norma Angélica Hernández-Pineda<sup>31</sup>, Jorge Granados-Kraußes<sup>31</sup>, María Adriana Rodríguez-Vázquez<sup>32</sup>, Delfino Torres-Valle<sup>33</sup>, Carlos Cortés-Reyes<sup>34</sup>, Francisco Medrano-López<sup>35</sup>, Jessica Arfeel Pérez-Gómez<sup>36</sup>, Annel Martínez-Ríos<sup>36</sup>, Antonio Aguilar-De los Santos<sup>37</sup>, Benenice Sarafín-Díaz<sup>38</sup>, Vilma Carolina Bekker-Méndez<sup>39</sup>, Minerva Mata-Rocha<sup>40</sup>, Blanca Angélica Morales-Castillo<sup>41</sup>, Omar Alejandro Sepúlveda-Robles<sup>42</sup>, Julián Ramírez-Bello<sup>43</sup>, Haydée Rosas-Vargas<sup>44</sup>, Alfredo Hidalgo-Miranda<sup>4</sup>, Juan Manuel Mejía-Arangurí<sup>2,45\*</sup> and Silvia Jiménez-Morales<sup>46</sup>

## OPEN ACCESS

### Edited by:

Alessandro Iacovi,  
Azienda Ospedaliera Civile di Rieti,  
Marche Nord, Italy

### Reviewed by:

Annalisa Lombi,  
University of Bologna, Italy  
Azz Mahmoud Kameel,  
Cairo University, Egypt

### \*Correspondence:

Juan Manuel Mejía-Arangurí  
arangurimj@hotmail.com;  
juan.mejia@imss.gob.mx  
Silvia Jiménez-Morales  
jimenezsilviam@imss.gob.mx

### Specialty section:

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

Received: 29 June 2020

Accepted: 17 August 2020

Published: 21 September 2020

<sup>1</sup> Departamento de Hemato-Oncología, Hospital Infantil de México Federico Gómez, Secretaría de Salud, México City, México; <sup>2</sup> Programa de Maestría y Doctorado en Ciencias Médicas de la Facultad de Medicina, Universidad Nacional Autónoma de México (UNAM), México City, México; <sup>3</sup> Unidad de Investigación Médica en Epidemiología Clínica, Unidad Médica de Alta Especialidad Hospital de Pediatría "Dr. Silvestre Frank Freund", Centro Médico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, México City, México; <sup>4</sup> Laboratorio de Genética del Cáncer, Instituto Nacional de Medicina Genómica (INMIGEN), México City, México; <sup>5</sup> Universidad Michoacana, Campus Tijuana, Tijuana, México; <sup>6</sup> Facultad de Estudios Superiores Zaragoza, Universidad Nacional Autónoma de México, México City, México; <sup>7</sup> Servicio de Hematología Pediátrica, Centro Médico Nacional "La Raza", Hospital General "Gaudencio González García", Instituto Mexicano del Seguro Social (IMSS), México City, México; <sup>8</sup> Genética Computacional, Instituto Nacional de Medicina Genómica (INMIGEN), México City, México; <sup>9</sup> Servicio de Hematología Pediátrica, Centro Médico Nacional "Siglo XXI", IMSS Hospital de Pediatría "Dr. Silvestre Frank Freund", Instituto Mexicano del Seguro Social (IMSS), México City, México; <sup>10</sup> Servicio de Cirugía Pediátrica, Hospital General "Gaudencio González García", Centro Médico Nacional Siglo XXI (CMN) "La Raza", Instituto Mexicano del Seguro Social (IMSS), México City, México; <sup>11</sup> Servicio de Hematología Pediátrica, Hospital General Regional "Carlos McGregor Sánchez Navarro", Instituto Mexicano del Seguro Social (IMSS), México City, México; <sup>12</sup> Clínica Pediátrica del Hospital Regional "General Ignacio Zaragoza", Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), México City, México; <sup>13</sup> Servicio de Onco-Pediatría, Hospital Juárez de México, Secretaría de Salud (SS), México City, México; <sup>14</sup> Servicio de Oncología, Hospital Pediátrico de Morelia, Secretaría de Salud de la Ciudad de México (SSCDM), México City, México; <sup>15</sup> Servicio de Pediatría, Hospital General de Toluca, Secretaría de Salud (SS), México City, México; <sup>16</sup> Servicio de Hematología Pediátrica, Hospital General de México, Secretaría de Salud (SS), México City, México;

<sup>11</sup> Coordinación Clínica y Servicio de Cirugía pediátrica, Hospital General Regional (HGR) No. 1 "Dr. Carlos Mac Gregor Sánchez Navarro", Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>12</sup> Servicio de Hematología Pediátrica, Centro Médico Nacional "20 de Noviembre", Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), Mexico City, Mexico; <sup>13</sup> Jefatura de Enseñanza, Hospital Pediátrico de Interoaxco, Secretaría de Salud de la Ciudad de México (SSCDM), Mexico City, Mexico; <sup>14</sup> Jefatura de Enseñanza, Hospital Pediátrico de Interoaxco, Secretaría de Salud de la Ciudad de México (SSCDM), Mexico City, Mexico; <sup>15</sup> Servicio de Pediatría, Hospital General Zona (HGZ) No. 3 "Dr. Gilberto Flores Arguente", Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>16</sup> Jefatura de Enseñanza, Hospital General de Ecatepec "Las Américas", Instituto de Salud del Estado de México (ISEM), Mexico City, Mexico; <sup>17</sup> Jefatura de Enseñanza, Hospital Pediátrico La Villa, Secretaría de Salud de la Ciudad de México (SSCDM), Mexico City, Mexico; <sup>18</sup> Jefatura de Enseñanza, Hospital Pediátrico San Juan de Aragón, Secretaría de Salud (SS), Mexico City, Mexico; <sup>19</sup> Servicio de Pediatría, Hospital General de Tacubaya, Secretaría de Salud de la Ciudad de México (SSCDM), Mexico City, Mexico; <sup>20</sup> Coordinación Clínica de Educación e Investigación en Salud, Hospital General de Zona (HGZ) No. 47, IMSS, Mexico City, Mexico; <sup>21</sup> Servicio de Cirugía Pediátrica, Hospital Pediátrico de Modurama, Secretaría de Salud de la Ciudad de México (SSCDM), Mexico City, Mexico; <sup>22</sup> Coordinación de Investigación en Salud, Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>23</sup> Pediatría, Hospital Materno-Pediátrico de Xochimilco, Secretaría de Salud de la Ciudad de México (SSCDM), Mexico City, Mexico; <sup>24</sup> Jefatura de Enseñanza, Hospital Pediátrico de Coyacoacán, Secretaría de Salud de la Ciudad de México (SSCDM), Mexico City, Mexico; <sup>25</sup> Coordinación Clínica y Pediatría del Hospital General de Zona 76, Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>26</sup> Coordinación Clínica y Pediatría del Hospital General de Zona 68, Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>27</sup> Coordinación Clínica y Pediatría del Hospital General de Zona 71, Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>28</sup> Pediatría, Hospital General Dr. Darío Fernández Fierro, Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), Mexico City, Mexico; <sup>29</sup> Coordinación Clínica y Servicio de Pediatría, Hospital General Regional (HGR) No. 72 "Dr. Vicente Santos Guajardo", Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>30</sup> Cirugía Pediátrica del Hospital Regional "General Ignacio Zaragoza", Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), Mexico City, Mexico; <sup>31</sup> Coordinación Clínica y Pediatría del Hospital General de Zona 66, Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>32</sup> Coordinación Clínica y Pediatría del Hospital General de Zona 57, Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>33</sup> Hospital de Infectología "Dr. Daniel Méndez Hernández", "La Rusa", Instituto Mexicano del Seguro Social (IMSS), Unidad de Investigación Médica en Inmunología e Infectología, Mexico City, Mexico; <sup>34</sup> Unidad de Investigación Médica en Genética Humana, UNAM-Hospital de Pediatría "Dr. Silvestre Ríos Pérez", Centro Médico Nacional "Siglo XXI", Instituto Mexicano del Seguro Social (IMSS), Mexico City, Mexico; <sup>35</sup> Unidad de Investigación, Hospital Juárez de México, Mexico City, Mexico

**Background:** Acute lymphoblastic leukemia (ALL) is the main type of cancer in children. In Mexico and other Hispanic populations, the incidence of this neoplasm is one of the highest reported worldwide. Functional polymorphisms of various enzymes involved in the metabolism of xenobiotics have been associated with an increased risk of developing ALL, and the risk is different by ethnicity. The aims of the present study were to identify whether *NQO1*, *CYP2E1*, and *NAT2* polymorphisms or some genotype-environmental interactions were associated with ALL risk in Mexican children.

**Methods:** We conducted a case-control study including 478 pediatric patients diagnosed with ALL and 284 controls (children without leukemia). Ancestry composition of a subset of cases and controls was assessed using 32 ancestry informative markers. Gene-environmental interactions for the exposure to hydrocarbons were assessed by logistic regression analysis.

**Results:** The polymorphisms rs1801280 (OR 1.54, 95% CI 1.21–1.93), rs1799929 (OR 1.96, 95% CI 1.55–2.49), and rs1208 (OR 1.44, 95% CI 1.14–1.81) were found to increase the risk of ALL; being the risks higher under a recessive model (OR 2.20, 95% CI 1.30–1.71, OR 3.87, 95% CI 2.20–6.60, and OR 2.26, 95% CI 1.32–3.87, respectively). Gene-environment interaction analysis showed that *NAT2* rs1799929 TT genotype confers high risk to ALL under exposure to fertilizers, insecticides, hydrocarbon derivatives, and parental tobacco smoking. No associations among *NQO1*, *CYP2E1*, and ALL were observed.

**Conclusion:** Our study provides evidence for the association between NAT2 polymorphisms/gene-environment interactions, and the risk of childhood ALL in Mexican children.

**Keywords:** acute lymphoblastic leukemia, ancestry informative markers, NAT2 polymorphisms, association study, Mexican mestizo

## INTRODUCTION

Acute lymphoblastic leukemia (ALL) represents the most frequent type of cancer in pediatric population worldwide. Additionally, Mexico and other Hispanic populations have one of the highest incidence and mortality rates for this cancer around the globe (1, 8). The etiology is unclear; however, it has been recognized that the interaction between genetic and environmental factors may play a role in the development of the disease. It has been suggested that enzymes involved in the metabolism of carcinogenic agents are associated with ALL susceptibility; as well, it could explain the differences in survival rates reported across populations (9–11).

Hydrocarbons, parental smoking, alcohol drinking, pesticides, and air pollution have been reported as the main type of xenobiotics related with an increased risk of developing ALL in children (12–16). Afterwards, these compounds enter into the body, they undergo biotransformation and elimination processes by xenobiotic metabolizing enzymes (2, 17, 18). Phase I enzymes are responsible for catalyzing reactions of hydroxylation, reduction and oxidation of xenobiotics, ultimately, transforming these into more toxic compounds. Phase II enzymes participate in conjugation processes, such as glucuronidation, acetylation, and methylation, converting the metabolites into non-reactive and water-soluble products which could be more easily eliminated from the body (18).

Cytochrome P450 Family 2 Subfamily E Member 1 (CYP2E1) are phase I enzymes, whereas, NAD(P)H quinone oxidoreductase 1 (NQO1) and Arylamine N-acetyltransferase 2 (NAT2) are phase II enzymes, which are commonly implicated in xenobiotics metabolism (2). It is well-recognized that genetic polymorphisms could influence the activity of these enzymes by affecting at a transcription level (CYP2E1), enzymes activity (NQO1) or at a protein level (NAT2) (2, 19). In addition, diverse studies have reported that polymorphisms in genes encoding xenobiotic metabolizing enzymes could increase the risk of developing childhood ALL, early relapse, chemotherapy-related toxicity, treatment resistance and have been associated with lower survival rates in patients with this type of cancer (2, 17). However, the results are inconsistent across populations suggesting the possibility that susceptibility alleles could be associated with ALL in a race-specific manner, as it has been reported for other ALL-predisposing genetic polymorphisms (20, 21).

The aims of the present study were to identify whether NQO1, CYP2E1, and NAT2 polymorphisms, or some genotype-environmental interactions, were associated with ALL risk in Mexican children. As well, to provide a picture of the genetic ancestry and admixture patterns of the patients with ALL from central Mexico.

## MATERIALS AND METHODS

### Subjects

A case-control study was conducted by the Mexican Interinstitutional Group for the Identification of the Causes of Childhood Leukemia (MIGCCL), from August 1, 2014, to July 31, 2016. Cases and controls were younger than 17 years, residents of the Metropolitan Area of Mexico City.

### Cases

Cases were recruited from eight public hospitals of Mexico City in which an estimated 69.9% of children with leukemia who reside in Mexico City are attended. Acute lymphoblastic leukemia diagnosis was established according to clinical features, and bone marrow aspirate findings: cell morphology, immunophenotype, and genetics, as defined by the 2008 WHO classification of lymphoid neoplasms. Children with Down syndrome were excluded from the analysis. Case registration required that trained personnel were assigned to each participating hospital to identify incident cases of leukemia through reviews of clinical charts. Afterwards, parents were approached and invited to participate.

The participating public hospitals and health institutions were: (1) Hospital de Pediatría, "Dr. Silvestre Frenk Freund" Centro Médico Nacional (CMN) "Siglo XXI", Instituto Mexicano del Seguro Social (IMSS); (2) Hospital Infantil de México Federico Gómez, Secretaría de Salud (SS); (3) Hospital General "Dr. Gaudencio González García", CMN "La Raza", IMSS; (4) Hospital General Regional "Dr. Carlos McGregor Sánchez Navarro", IMSS; (5) Hospital Juárez de México, SS; (6) Hospital Pediátrico de Mectezuma, Secretaría de Salud de la Ciudad de México (SSCDMX); (7) Hospital General de México, SS; and (8) Hospital CMN "20 de Noviembre", Instituto de Seguridad Social al Servicio de los Trabajadores del Estado (ISSSTE). For this study the sample size was calculated with 250 cases and 250 controls, this being sufficient for obtaining a power of 80% to identify the associations. In parallel to this study the inclusion of cases continued for assessing the prognostic value of the NQO1, CYP2E1, and NAT2 polymorphisms studied here on ALL survival. Those cases were also included in the present work for increasing the precision of the assessed interactions.

### Controls

The controls were selected from second-level hospitals of the same health institution (IMSS, SS, SSCDMX, and ISSSTE) that referred the children with ALL to the third-level care hospitals. The controls were children without leukemia who were treated at different hospital departments, such as ambulatory surgery, pediatrics, ophthalmology and orthopedic outpatient

clinics, and the emergency room. Children with diagnoses of neoplasms, hematological diseases, allergies, infections, and congenital malformations were not selected as controls. The interviewers were provided with a list of characteristics of individual leukemia cases, and charged with identifying one control per case according to sex, age ( $\pm 18$  months), and health institution. When more than one control per case met the matching criteria, the control that was closest in age to the case was selected. If the controls were of the same age, one control was selected randomly (by tossing a coin). Moreover, if no control of the same sex was found after three visits to the same hospital, a control matched by age ( $\pm 18$  months) was selected.

Trained personnel were assigned to each of the second-level hospitals to search for controls who fulfilled the selection criteria. When a control was identified, parents were invited to participate.

The Ethics and National Committee of Scientific Research approved this study with number R-2013-783-062. Additionally, we obtained approval by the Institutional Scientific Review Committees of each health institution to carry out the study.

## Study Overview

The interview was carried out using the same questionnaire for cases and controls. This questionnaire was previously standardized and adapted from the questionnaire module of the National Cancer Institute. General information obtained from the interviews included the child's history, socioeconomic status (SES), demographic characteristics, and the parents information (age at pregnancy, educational level, etc.).

Study variables were child's sex, age, and birthweight ( $< 3,500$  and  $\geq 3,500$  grams), family history of cancer, maternal years of education, after birth home exposure to fertilizers, insecticides, and other hydrocarbon derivatives (benzene, solvents, glues, petroleum products, etc.), and parental tobacco smoking.

For the exposure of the parents to tobacco smoking, the information was obtained from the in-person interview and it was assessed in the four exposure periods: (1) pre-conceptionally, (2) during pregnancy, (3) during lactation, and (4) the last year before leukemia diagnosis or interview in cases and controls. The number of cigarettes smoked during each exposure period was calculated. Parental smoking was classified into three groups: (1) when both parents were lifelong non-smokers or ex-smokers, who stopped smoking more than 1 year before the birth of the index child, (2) when either parent smoked  $< 5$  cigarettes a day, and (3) when at least one parent smoked 6 or more cigarettes daily. Passive smoking in the child was considered when at least one of the parents smoked in the presence of the child for at least three times a week during the year before diagnosis/interview in cases and controls, respectively.

Parental years of education were used as an indicator of SES as it has been previously used by the Childhood Leukemia International Consortium (CLIC) (0–9 years, 9.1–12.9 years [reference category],  $\geq 13$  years of education).

## DNA Extraction

Genomic DNA from saliva or peripheral blood was obtained according to the ORAGENE Purification Kit (DNA

Genotek Inc., ON, Canada) and the Genra Kit (Genra Systems Inc, Minneapolis, MN) manufacturer's instructions, respectively. DNA purity and concentration were determined by spectrophotometry (Nanodrop-1000).

## Genotyping of NQO1, CYP2E1, and NAT2 SNPs

Genotyping of rs1800566 (NQO1), rs3813867 (CYP2E1) and rs1041983 (C282T), rs1801280 (T341C), rs1799929 (C481T), rs1799930 (G590A), rs1208 (A803G), and rs1799931 (G857A) from NAT2 SNPs was performed using the 5' exonuclease technique and TaqMan MGB chemistry in a QuantStudio 5 system according to the manufacturer's instructions (ThermoFisher, Foster City, CA). TaqMan probes used were C\_\_2091255\_30 (rs1800566), C\_\_2431875\_10 (rs3813867), C\_\_8684065\_20 (rs3813867), C\_\_1204093\_20 (rs1801280), C\_\_1204092\_20 (rs1799929), C\_\_1204091\_10 (rs1799930), C\_\_572569\_20 (rs1208). Each polymerase chain reaction (PCR) reaction contained 25 ng of genomic DNA, 2.5  $\mu$ l of TaqMan master mix, 0.0625  $\mu$ l of 40X assay mix, and ddH<sub>2</sub>O up to a final volume of 5  $\mu$ l. The PCR protocol included denaturing at 95°C for 10 min, followed by 40 cycles of denaturing at 95°C for 15 s and annealing and extension at 60°C for 1 min. The genotype of each sample was assigned automatically by measuring the allele-specific fluorescence using QuantStudio Design and Analysis software 5 for allelic discrimination (Applied Biosystems, Foster City, CA). The overall genotype call rate was 99.8%, and a subset of randomized samples were genotyped in duplicate and showed 100% reproducibility.

## Ancestry Informative Markers Selection and Ancestry Composition Analysis

To know the ancestry composition of the ALL pediatric cases, and to account for the effects of population stratification in our case-control association study, we analyzed a subset of ALL cases and controls using a panel of 32 single nucleotide polymorphisms (SNPs) as ancestry informative markers (AIMs). The SNPs chosen for inclusion were based on AIMs sets that has been validated in a large group of Mexican subjects. This panel included 1 to 2 SNPs in all chromosomes, with the exception of chromosome 18 (22). ALL cases and controls were paired by gender, age, and living area in Mexico-City. To estimate individual and global ancestry we used the STRUCTURE software considering the European and Native-American (NA) populations as the two main contributors to the largest Mexican ethnic group (Mexican-Mestizo) (3). Principal component analysis (PCA) was used to infer population structure within our samples (23).

## Association and Gene-Gene Interaction Analyses

The polymorphisms NQO1 (rs1800566), CYP2E1 (rs3813867), and NAT2 [rs1041983 (C282T), rs1801280 (T341C), rs1799929 (C481T), rs1799930 (G590A), rs1208 (A803G), and rs1799931 (G857A)] were selected based on previous results in leukemia where they were the most consistently associated with leukemia

risk (2, 11, 24–26). Association among alleles/genotypes and ALL was tested by comparing cases and controls. There was 99% concordance of randomly repeated samples during the genotyping. Bonferroni correction test was applied to data, then, the *p*-value for statistical significance varied according to the number of parameters. In addition, all SNPs were evaluated under the co-dominant, dominant, and recessive genetic models using the FINETTI program (<http://ihg.gsf.de/cgi-bin/hw/hwa1.pl>). Linkage disequilibrium (LD) and haplotyping of NAT2 SNPs were assessed by using the haploview software (27). The Multifactor Dimensionality Reduction (MDR) interaction model was generated by multifactor dimensionality reduction software (V 3.0.2) (28). Results were considered to be significant at  $p \leq 0.05^*$ .

### NAT2 Acetylator Frequencies Analysis

Patients were phenotyped as rapid acetylators and slow acetylators. Haplotypes containing the wild-type (WT) NAT2 alleles (rs1041983 (C282T), rs1801280 (T341C), rs1789929 (C481T), rs1799930 (G590A), rs1208 (A803G), and rs1799931), named NAT2<sup>4</sup>, were classified as a rapid acetylator phenotype. Haplotypes containing more than one mutant allele were identified as slow acetylators and the remainder as an intermediate acetylator phenotype, using the NAT2PRED web-server (<http://nat2pred.rit.albany.edu>) (29). NAT2 alleles were identified by using [http://nat.mbg.duth.gr/Human%20NAT2%20alleles\\_2013.htm](http://nat.mbg.duth.gr/Human%20NAT2%20alleles_2013.htm).

### Evaluation of Interactions Among Polymorphisms and Exposure to Hydrocarbons

To search for gene-environment interactions we used the MDR approach, since this program incorporates information from several loci and environmental factors to identify combinations of both factors that are associated with the risk of a disease (30).

### Statistical Analysis

Analyses were performed using SPSS version 21 (IBM Statistical Package for the Social Sciences, Inc., Chicago, IL, USA). Descriptive analyses were conducted. Odds ratio (OR) and 95% confidence intervals (95% CI) were estimated by unconditional logistic regression. The logistic regression model was constructed as follows: (1) the overall interaction in the model was evaluated, (2) it was assessed whether study variables had a global confounding effect, those variables that, in the stratified analysis the ORs prior to (crude OR) and after stratifying (aOR) showed a difference between the two ORs >10%. These variables alongside with the matching variables (child's sex, age, and health institution), were included in the multivariate analyses. As a result, the most parsimonious model included: child's sex, age, paternal education level, maternal education level, maternal age at pregnancy, active smoking of the mother after birth, alcoholism of the mother before pregnancy, family cancer history, health institution, birthweight, maternal exposure to hydrocarbons at home during pregnancy, X-rays exposure during pregnancy, alcoholism of the father before pregnancy, and

active smoking of the father after birth. Accordingly, the adjusted ORs (aOR) were calculated.

Hardy-Weinberg Equilibrium (HWE) test was performed using the FINETTI program (<http://ihg.gsf.de/cgi-bin/hw/hwa1.pl>). Alleles and genotypes frequencies were compared among groups by using chi-square and Fisher's exact test, when appropriate, which are implemented in the STATCALC program (Epi Info v.6.02 software, Centers for Disease Control and Prevention, Atlanta, GA). The level of significance was set at 5%. To evaluate ancestry composition of our sample, STRUCTURE software was used assuming two populations (European and NA), and each analysis was performed at least three times using >100,000 replicates and 20,000 burn-in cycles under admixed model. To determine the statistical power of our study, we used Quanto software (<http://hydra.usc.edu/gxe>) accounting minor allele frequency (MAF) of all SNPs in the control the group; likewise, considering a recessive genetic model, and odd ratio (OR) of 2.0, the prevalence of ALL in Mexican children, and the sample size.

## RESULTS

### Patient Demographic Data

A total of 469 cases were diagnosed with ALL in participant hospitals during the study period. In addition, a total of 285 controls were recruited. Demographic, clinical features and exposure data of the study population are displayed in **Supplementary Table 1**. The following variables: child's age, parental education level, parental smoking, alcohol drinking by the father, and maternal exposure to hydrocarbons at home showed statistically significant differences between the total of cases with ALL and controls. Cases had a less proportion of active smoking by the mother after birth, alcohol drinking by both parents before pregnancy, and maternal exposure to hydrocarbons at home before pregnancy, but higher proportion of maternal exposure to hydrocarbons at home during pregnancy than controls (**Supplementary Table 1**). The associations when 279 cases were analyzed and then, when all the available 469 cases were included, were very similar (**Supplementary Table 2**). It was therefore decided to include the total number of cases in subsequent analyses.

### NAT2 Polymorphisms Are Associated With Acute Lymphoblastic Leukemia

Except for the rs1041983 (C282T) and rs1799931 (G857A), the remaining SNPs were found in HWE in the control population. The association analysis between individual SNPs and ALL are described in **Table 1**. We documented statistically significant differences in the distribution of rs1801280 (T341C) (OR 1.54, 95% CI 1.12–1.95), rs1789929 (C481T) (OR 1.95, 95% CI 1.53–2.47), and rs1208 (A803G) (OR 1.47, 95% CI 1.175–1.86) NAT2 SNPs.

In the genotype analysis, an association with ALL was observed, both under heterozygous and recessive models to the rs1801280 (T341C) (OR 1.44, 95% CI 1.04–1.98, and OR 2.25, 95% CI 1.33–3.80, respectively) and rs1789929

**TABLE 1 |** Genotype and allele frequency of NAT2, NQO1, and CYP2E1 polymorphisms in children with acute lymphoblastic leukemia and controls.

Gene	SNP	Genotypes/Alleles	Cases 466 n (%)	Controls 395 n (%)	OR (95% CI)	aOR (95% CI)
NAT2*	rs1041983 (C282T)	CC (ref)	270 (57.7)	145 (36.6)	—	—
		CT	152 (32.5)	102 (25.8)	0.80 (0.59–1.10)	0.73 (0.49–1.09)
		TT	46 (9.8)	25 (6.4)	0.71 (0.43–1.14)	0.56 (0.30–1.07)
		C	692 (73.9)	392 (99.5)	—	—
	rs1801280 (T341C)	T (ref)	244 (52.1)	172 (43.3)	0.80 (0.60–1.04)	—
		TT (ref)	220 (47.0)	165 (41.5)	—	—
		TC	182 (38.8)	95 (23.7)	<b>1.44 (1.04–1.98)</b>	<b>1.70 (1.14–2.53)</b>
		CC	66 (14.1)	22 (5.6)	<b>2.25 (1.32–3.80)</b>	<b>2.60 (1.08–6.60)</b>
	rs1799929 (C481T)	T	622 (76.5/0)	425 (75.6)	—	—
		C	214 (23.5/0)	139 (34.9)	<b>1.54 (1.12–1.95)</b>	—
		CC (ref)	203 (43.4)	167 (42.4)	—	—
		CT	186 (39.7)	90 (22.7)	<b>1.59 (1.16–2.20)</b>	<b>1.80 (1.30–2.60)</b>
	rs1799930 (G590A)	TT	79 (16.9)	18 (4.4)	<b>3.61 (2.08–6.27)</b>	<b>3.72 (1.91–7.25)</b>
		C	592 (63.2)	430 (76.5)	—	—
		T	344 (36.7)	132 (23.3)	<b>1.95 (1.52–2.47)</b>	—
		GG (ref)	286 (62.7)	232 (78.6)	—	—
	rs1208 (A803G)	GA	72 (15.4)	60 (15.2)	0.69 (0.47–1.01)	0.70 (0.49–1.00)
		AA	9 (1.9)	1 (0.4)	5.18 (0.65–41.12)	8.72 (0.95–80.14)
		G	844 (90.4)	504 (99.0)	—	—
		A	90 (9.6)	62 (11.0)	0.86 (0.61–1.22)	—
rs1799931 (G857A)	AA (ref)	203 (43.6)	149 (37.3)	—	—	
	AG	195 (41.8)	110 (27.8)	1.30 (0.95–1.78)	1.44 (0.97–2.14)	
	GG	68 (14.6)	21 (5.2)	<b>2.38 (1.39–4.05)</b>	<b>3.25 (1.73–6.10)</b>	
	A	601 (64.5)	408 (72.9)	—	—	
rs1799931 (G857A)	G	321 (35.4)	152 (27.1)	<b>1.47 (1.75–1.88)</b>	—	
	GG (ref)	222 (69.2)	191 (68.5)	—	—	
	GA	115 (24.7)	58 (20.1)	1.22 (0.84–1.78)	1.37 (0.81–1.98)	
	AA	28 (6.0)	22 (7.1)	<b>0.52 (0.30–0.89)</b>	<b>0.34 (0.17–0.71)</b>	
NQO1**	rs3813867	G	759 (91.6)	438 (78.5)	—	—
		A	171 (18.4)	120 (21.5)	0.82 (0.62–1.09)	—
		CC (ref)	154 (33.0)	78 (28.0)	—	—
		CT	224 (48.0)	145 (52.0)	0.78 (0.55–1.10)	0.80 (0.52–1.22)
		TT	89 (19.1)	50 (20.1)	0.80 (0.52–1.24)	0.91 (0.52–1.58)
CYP2E1**	rs3813867	C	322 (57.0)	301 (54.0)	—	—
		T	402 (63.0)	257 (68.1)	0.98 (0.71–1.09)	—
		GG (ref)	292 (64.3)	182 (66.2)	—	—
		GC	150 (32.0)	89 (32.4)	1.05 (0.76–1.45)	0.91 (0.60–1.38)
		CC	12 (2.6)	4 (1.5)	<b>1.87 (0.59–5.88)</b>	<b>1.39 (0.34–5.70)</b>
		G	734 (80.8)	453 (82.4)	—	—
		C	174 (19.2)	97 (17.6)	1.10 (0.84–1.45)	—

OR, crude odds ratio; aOR, adjusted odds ratio by child's sex, age, paternal education level, maternal education level, maternal age at pregnancy, Active smoking of the mother after birth, alcoholism of the mother before pregnancy, family cancer history, health insurance, birthweight, Maternal exposure to hydrocarbons at home during pregnancy,  $\alpha$ -rays exposure during pregnancy, alcoholism of the father before pregnancy, Active smoking of the father after birth. Genotyping rate \* = 96% and \*\* = 98.6%. Bold values represent the most statistically significance results.

(C481T) (OR 1.59, 95% CI 1.16–2.20 and OR 3.61, 95% CI 2.08–6.27, respectively); as well, under recessive model for the rs1208 (A803G) (OR 2.38, 95% CI 1.39–4.05). CYP2E1 (rs3813867) and NQO1 (rs1800566) SNPs were not associated with ALL (Table 1).

No interactions among NQO1 (rs1800566), CYP2E1 (rs3813867), and NAT2 (rs1799930, G590A) and the exposure variables were noted (Supplementary Table 3). Nevertheless, homozygotes genotypes TT and AA of NAT2 SNPs rs1041983 (C282T) and rs1799931 (G857A), respectively, were associated

with a reduced risk of ALL under exposure to hydrocarbons by the mother at home before and during pregnancy and after birth (HEMAB), drug consumption, active smoking, alcohol consumption and insecticide exposure by the mother before pregnancy, whilst being pregnant and while child feeding. As well as active smoking and alcohol drinking by father before and during pregnancy. NAT2 variants rs1799929 and rs1208 variants, which were associated with ALL risk, exhibited an increased ORs under maternal hydrocarbons exposure and drug consumption. In the same way, active smoking, and alcohol consumption by the parents during pregnancy incremented the risk to ALL in homozygotes cases for mutant alleles to both SNPs. In fact, the rs1799929 (C481T) and active smoking by the father after birth displayed the highest OR (4.49, CI 95% 2.46–8.17) to ALL observed in the present analysis. The rs1799930 (NAT2), rs1800566 (NQO1), and rs3813867 (CYP2E1) were not associated neither after adjusting by these variables (Supplementary Table 3).

### Linkage Disequilibrium (LD) and Haplotypes Analyses

The assessment of LD displayed low LD among NAT2 polymorphisms (Supplementary Figure 1). By including all NAT2 variants, we found 11 haplotypes showing frequencies >0.01 as described in Table 2. Rapid alleles as NAT2\*4, NAT2\*11A, NAT2\*12A, NAT2\*12C, and NAT2\*13A were identified. Slow alleles as NAT2\*5B, NAT2\*5C, NAT2\*5V, NAT2\*6A, NAT2\*7A, and NAT2\*7B, were detected. Statistically significant differences were noted for NAT2\*4, NAT2\*7B, and NAT2\*13A distribution between cases and controls. Even though NAT2\*5C, NAT2\*11A, and NAT2\*12C, also showed significant differences between cases and controls. Nonetheless, due to the CI values, the associations should be taking with caution (Table 2).

### Frequency of Rapid and Slow Acetylators

Patients harboring wild type for all of the NAT2 polymorphisms were classified as rapid acetylators. Those with homozygous mutants or heterozygous for more than one of the polymorphisms were phenotyped as slow acetylators, being the remainder considered as intermediate acetylators (Table 3). Overall cases and controls, 458 and 269 phenotypes, respectively, were established. Rapid acetylator frequency was 20.7%, intermediate phenotype was observed in 45.9% and slow in 33.4% of the children with ALL. Meanwhile, 25.3% of the control population was comprised by rapid acetylators, and intermediate and slow were found in 39.8 and 34.9%, respectively. No statistically significant differences were found (Table 3).

### Gene-Gene Interaction Between NQO1-CYP2E1-NAT2 Polymorphisms

To investigate for interactions among NQO1-CYP2E1-NAT2 genotypes, we performed a MDR analysis including cases and controls having complete genotyping data. The model with the lowest prediction error and highest cross-validation

consistency (CVC) was selected (Table 4). The rs1799929 was the best factor model which showed statistically significant difference with testing accuracy (TBA) of 0.5856, a 10/10 CVC (Table 4) and contributing with 3.03% to the risk of ALL. The multilocus model with CVC (10/10) and minimum prediction error (maximum testing accuracy is 0.6337) was a three-factors model including the NAT2\_rs1801280, rs1799929, and rs1208. These three NAT2 SNPs interact together to collectively increase the risk to develop ALL (OR 6.59, CI 95% 4.05–10.71) (Figure 1, Table 4).

The interaction entropy analysis showed that NAT2-rs1799929 had the larger effect on the susceptibility to develop ALL (3.0%). A synergistic interaction (red lines) between NAT2-rs104183-rs1799931 and NAT2-rs1799929-rs1801280 SNPs, which reveals an epistasis effect between them, was observed. Redundancy (green lines) was observed between several NAT2 and CYP2E1 SNPs which means that jointly those SNPs provide less information than studying one SNP. Gold lines showed independence among SNPs (Figure 2).

### Genetic Structure of Mexican Children With Acute Lymphoblastic Leukemia and Controls

The ancestry composition of the first subset of patients with ALL ( $n = 166$ ) and controls ( $n = 167$ ) who met the selection criteria during the study period was evaluated. A total of 32 AIMs, previously validated in a Mexican population were used (22). We found a similar genetic background among cases and controls, which were enriched by a Native-American (NA) contribution (Figure 3).

### Gene-Environment Interaction in Acute Lymphoblastic Leukemia

MDR was used to explore gene-environment interactions. The following maternal exposure to hydrocarbons in the pre-conceptional, during pregnancy and after birth periods, were considered in the analysis: global exposure to hydrocarbons, active and passive smoking, alcohol, drugs and medicine consumption, wood smoke, insecticides, benzene, gasoline, and petroleum exposures. Whilst environmental exposure factors assessed for the father were: active and passive smoking and alcohol drinking mainly for the pre-conceptional and during pregnancy periods.

For maternal exposure to hydrocarbons during pre-conceptional and after birth periods, the best combination was determined for hydrocarbons exposure, active and passive smoking, drugs and medicine consumption and alcohol, as well as insecticides, benzene, gasoline, petroleum, and wood smoke exposures; in addition to active and passive smoking and alcohol consumption by the father.

Our analysis suggests that passive smoking by the father before conception (PSFBC) and during pregnancy (PSFDP) was the best factor with statistical significance (TBA 0.7087 CVC 10/10 and TBA 0.6868 CVC 10/10, respectively). The multifactor model with a minimum prediction error (TBA 0.7648)

**TABLE 2 |** Distribution of the most frequent NAT2 alleles in healthy children and in patients with acute lymphoblastic leukemia.

Haplotype	Alleles	Phenotype <sup>a</sup>	Cases (n = 925) n (%)	Controls (n = 268) n (%)	OR (95% CI)	P-values
CTOGAG	NAT2*4	Rapid	291 (31.3)	210 (78.9)	0.78 (0.62–0.97)	0.0283
OCTGGG	NAT2*28	Slow	193 (20.8)	105 (39.2)	1.16 (0.89–1.52)	0.3054
TTCGAA	NAT2*20	Slow	94 (10.1)	82 (30.2)	0.67 (0.49–0.92)	0.0109
TTCGAG	NAT2*16A	Slow	60 (6.4)	48 (17.6)	0.75 (0.49–1.13)	0.128
CTGGGG	NAT2*13A	Rapid	55 (5.9)	23 (8.6)	1.49 (0.89–2.52)	0.1075
TTCGAG	NAT2*13A	Rapid	17 (1.8)	23 (8.6)	0.44 (0.22–0.87)	0.0128
CTOGAA	NAT2*2A	Slow	16 (1.7)	19 (7.1)	0.51 (0.25–1.04)	0.0465
OCGGGG	NAT2*2C	Slow	28 (3.0)	6 (2.2)	2.91 (1.34–7.88)	0.007
CTTGAG	NAT2*11A	Rapid	28 (3.0)	1 (0.4)	17 (0.55–248.8)	<b>0.00001*</b>
CTTGGG	NAT2*12C	Rapid	19 (2.1)	2 (0.7)	5.91 (1.22–28.82)	<b>0.0095*</b>
TCTGAG	NAT2*2V	Slow	16 (1.7)	1 (0.4)	9.94 (1.22–201.58)	<b>0.0097*</b>

Haplotypes of NAT2 SNPs rs1041983-C/T, rs180566-T/C, rs1799929-C/T, rs1799929-G/A, rs12084V/G, and rs1799929-G/A were obtained by HAPLOVIEW. The haplotype frequencies <0.01 are not displayed.

<sup>a</sup>Phenotype inferred based on the belonging group NAT2-S. OR, crude odds ratio; 95% CI, 95% Confidence interval.

\*P-values remain under 10,000 permutation test. Bold values represent the most statistically significance results.

**TABLE 3 |** Association between acetylator phenotypes and risk of childhood acute lymphoblastic leukemia.

Acetylator phenotype	Cases n = 489 n (%)	Controls n = 265 n (%)	OR (95% CI)	aOR (95% CI)
Rapid (ref.)	95 (20.7)	68 (25.3)	—	—
Intermediate	210 (45.9)	107 (39.8)	1.40 (0.95–2.07)	1.83 (1.11–3.02)
Slow	153 (32.4)	94 (34.9)	1.16 (0.78–1.74)	1.21 (0.72–2.04)

OR, crude odds ratio; aOR, adjusted odds ratio by child's sex, age (<6 years), maternal education level, maternal age at pregnancy, active smoking of the mother after birth, alcoholism of the mother before pregnancy, family cancer history, health institution, birthweight, maternal exposure to hydrocarbons at home during pregnancy, x-ray exposure during pregnancy, alcoholism of the father before pregnancy, active smoking of the father after birth. Missing data for phenotype classification: 2 and 5.6% for cases and controls, respectively.

and CVC (10/10) was NAT2\_rs1799929-NQO1\_rs2811566-CYP2E1\_rs180566-alcohol consumption by the mother before pregnancy (ACMBP)-medicine consumption by the mother before pregnancy (MCMBP)-PSFBC suggesting that these factors jointly contributed to the etiology of ALL (Figure 4, Table 5).

After child's birth, the MDR identified to NAT2\_rs1799929 as the best one factor (TBA 6021 CVC 10/10) and the six-factor interaction model NAT2\_rs1041983, NAT2\_rs1799929, NAT2\_rs1799931, NQO1\_rs1800566-insecticide exposition after child's birth (IEAB)-active smoking by the father after child's birth (ASFAB) as the best model for ALL with a TBA of 0.6062 and CVC 9/10 (Table 5).

The interaction entropy analysis showed that PSFBC (13.39%), ACMBP (9.88), and MCMBP (9.65%) have the larger effect on susceptibility to ALL before pregnancy (Supplementary Figure 2), whilst the NAT2\_rs1799929 SNP was during pregnancy (4.15%) and after child's birth (4.21%) (Supplementary Figures 2, 3).

## DISCUSSION

Acute lymphoblastic leukemia, the most prevalent cancer in children, like other hematologic malignancies is likely to develop from a complex interaction between genetic and environmental factors. Hydrocarbons, such as benzene, pesticides, and air pollutants, are among the common xenobiotics that have been implicated in ALL etiology (13, 15). Detoxification enzymes play an important role to metabolize environmental carcinogens and it is well-known that polymorphisms in genes encoding for these enzymes may explain inter-individual differences in leukemia risk (16). To gain more knowledge on how these factors could be interplaying to influence disease risk and outcomes in ALL, and to provide a global picture of the genetic ancestry and admixture patterns of Mexican pediatric patients with ALL from central Mexico we conducted this study.

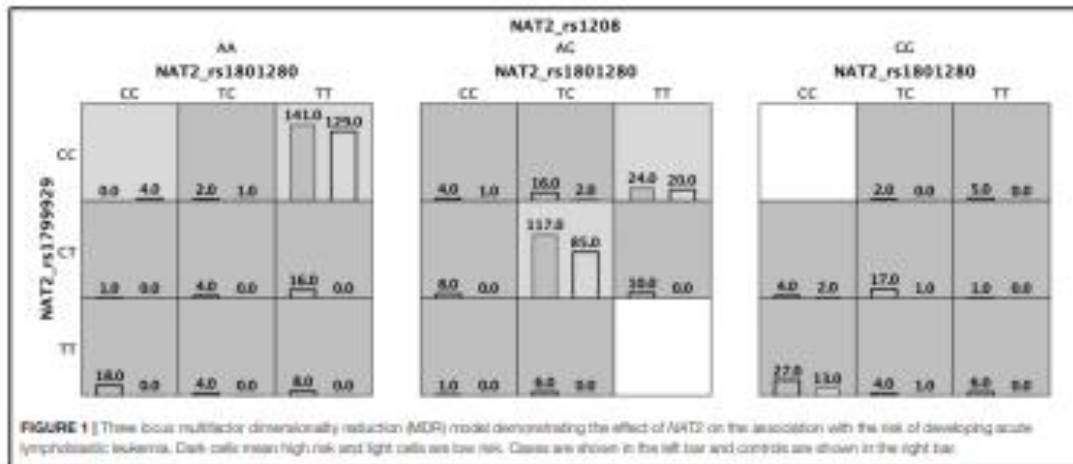
The present report focused on the effects of genetic variation in NQO1, CYP2E1A, and NAT2 polymorphisms and their interaction with environmental factors, such as home exposure to fertilizers, insecticides, hydrocarbon derivatives (benzene, petroleum products, etc.), and parental tobacco smoking and wood smoking, among other variables in the risk of developing childhood ALL. In the present study, we found that NAT2 polymorphisms, but not NQO1 and CYP2E1 polymorphisms were associated with ALL risk. NQO1 modifies internal exposure to bioactivated carcinogens. NQO1 has been previously described as an anticancer enzyme. There is a documented a relation between NQO1 (rs1800566) with several types of cancer, including infant leukemia (31–35); however, its association with ALL is controversial. Our results are in contrast with findings of some studies (34, 35), but they are consistent with other reports that did not show a significant association between NQO1 rs1800566 polymorphism and the risk of childhood ALL, when specific ethnic populations were analyzed (32, 36). Concerning CYP2E1, a fundamental contributor of the metabolism of low molecular weight compounds as ethanol, and a bioactivator of



**TABLE 4 |** Gene-gene interaction analysis between NAT2, NQO1, and CYP2E1 polymorphisms and acute lymphoblastic leukemia risk.

Number of factors	Model	TA	TBA	CVC	OR (CI 95%)	p-value*
1	<b>NAT2-rs1799929</b>	<b>0.585</b>	<b>0.595</b>	<b>10/10</b>	<b>1.90 (1.46-2.31)</b>	<0.0001
2	NAT2-rs1801280, NAT2-rs1799929	0.614	0.621	8/10	3.42 (3.26-3.57)	<0.0001
3	<b>NAT2-rs1801280, NAT2-rs1799929, NAT2-rs1208</b> **	<b>0.644</b>	<b>0.637</b>	<b>10/10</b>	<b>6.29 (4.05, 10.71)</b>	<b>&lt;0.0001</b>
4	NAT2-rs1041983, NAT2-rs1801280, NAT2-rs1799929, NAT2-rs1799921	0.602	0.598	6/10	3.75 (3.71-3.19)	<0.0001
5	NAT2-rs1041983, NAT2-rs1799929, NAT2-rs1799921, NQO1-rs1800566, CYP2E1-rs3813867	0.689	0.598	3/10	8.19 (4.19-6.15)	<0.0001
6	NAT2-rs1041983, NAT2-rs1799929, NAT2-rs1208, NAT2-rs1799921, NQO1-rs1800566, CYP2E1-rs3813867	0.715	0.549	5/10	3.89 (4.21-3.39)	<0.0001
7	NAT2-rs1041983, NAT2-rs1799929, NAT2-rs1799929, NAT2-rs1208, NAT2-rs1799921, NQO1-rs1800566, CYP2E1-rs3813867	0.739	0.575	4/10	7.23 (5.19-10.26)	<0.0001
8	NAT2-rs1041983, NAT2-rs1801280, NAT2-rs1799929, NAT2-rs1799929, NAT2-rs1208, NAT2-rs1799921, NQO1-rs1800566, CYP2E1-rs3813867	0.754	0.578	10/10	8.85 (5.23-13.59)	<0.0001

Best one-to-eight locus model of the NAT2-NQO1-CYP2E1 gene-gene interactions with cross-validation consistency (CVC) and prediction error per the *n*-locus model that have been obtained by the multifactor-dimensionality-reduction (MDR) of our data set. TA, training balance accuracy; TBA, testing balance accuracy; \*\*The best one-factor model and \*\*\*The best multilocus model. \*P-values from whole statistics data set. Bold values indicate the best single and model.



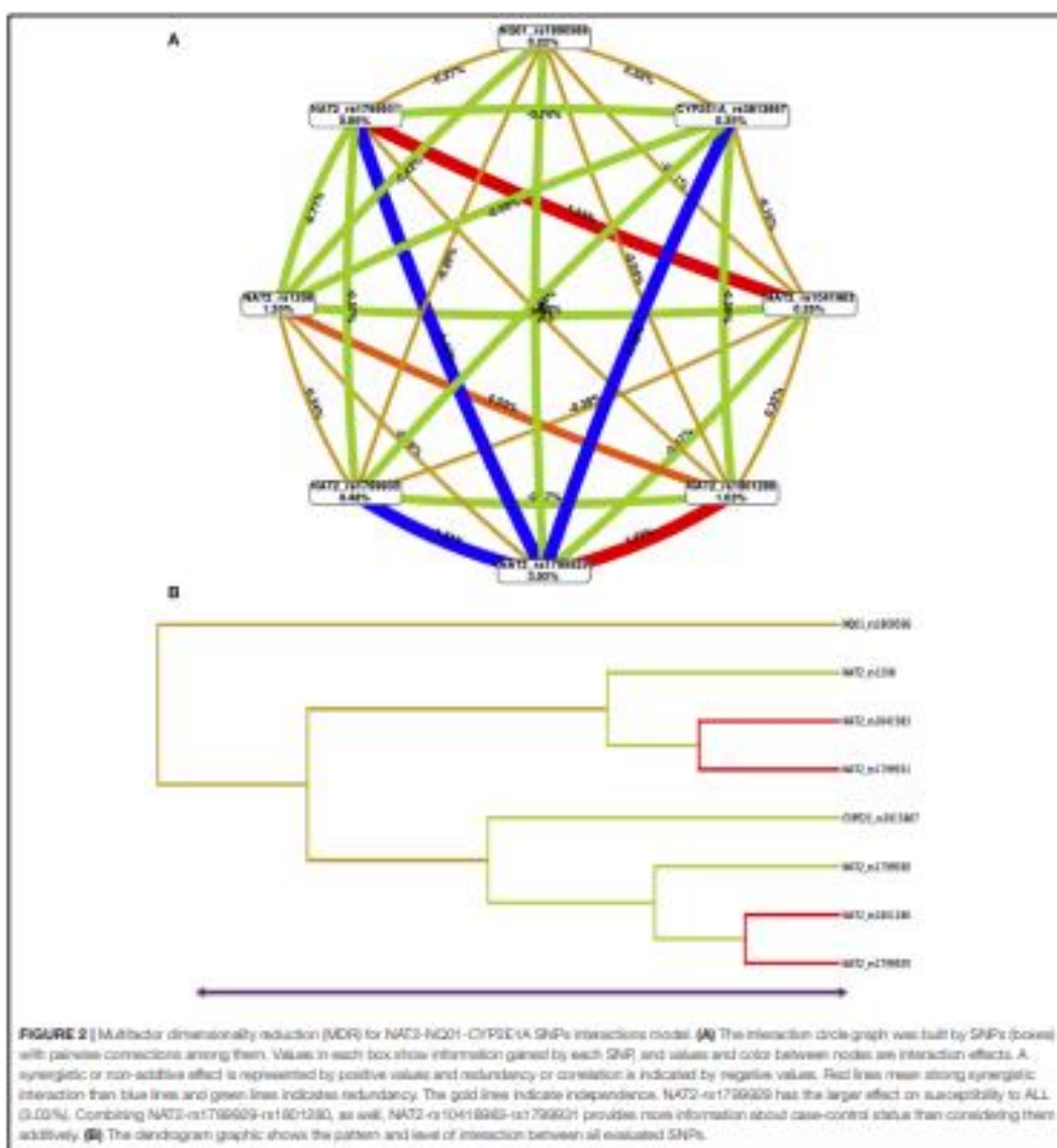
**FIGURE 1 |** Three locus multifactor dimensionality reduction (MDR) model demonstrating the effect of NAT2 on the association with the risk of developing acute lymphoblastic leukemia. Dark cells mean high risk and light cells are low risk. Cases are shown in the left bar and controls are shown in the right bar.

many procarcinogens including benzene, has been associated with an increased risk of ALL (26). A differential contribution by the variants among populations could be influenced by age and ethnicity (26, 35, 36). We have only studied one SNP for each gene so far. Since a number of functional polymorphisms of NQO1 and CYP2E1 have been identified, we cannot discard the contribution of these genes in the risk of ALL in Mexican children.

NAT2 is highly polymorphic and its activity is largely determined by coding single SNPs. To date, about 108 NAT2 alleles have been identified by the Gene Nomenclature Committee (37), which defines three metabolizer groups: slow, intermediate and rapid acetylators. The wild-type NAT2\*4 allele encodes a protein with high *N*-acetylation activity conferring a rapid acetylator phenotype (37, 38). Haplotypes containing more than one mutant allele at rs1041983, rs1801280 (T341C),

rs1799929 (C481T), rs1799930, rs1208 (A803G), and rs1799931 (G857A) defined the low acetylator phenotype (29, 37). Our data showed that rs1801280 (T341C), rs1799929 (C481T), rs1208 (A803G), and rs1799931 (G857A) polymorphisms confer higher risk of ALL in Mexican children, which increases in homozygote mutant carriers, compared with heterozygotes. Notably, these risks were even higher after adjusting by age, sex, and other environmental variables (Table 3).

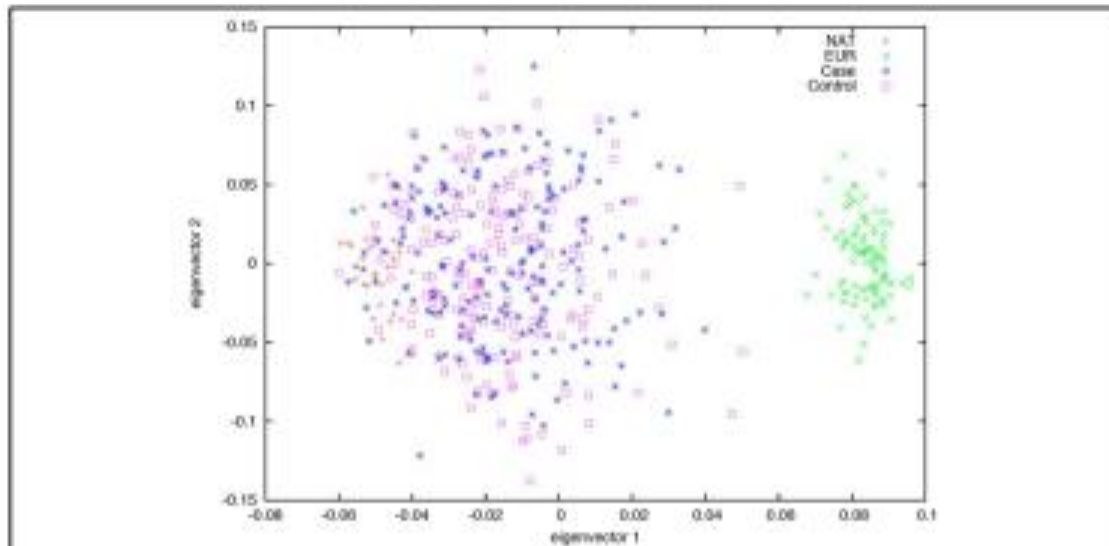
NAT2\*4, considered the most common allele involved in rapid acetylation, was frequent in our control group and statistically significant differences were observed between cases and controls. Consistent with our findings, NAT2\*5B has been reported to be the most common slow acetylator haplotype in Caucasian populations (4, 38). Notwithstanding, after multiple comparison tests, the statistical significance remained only for NAT2\*11A, NAT2\*12C, and NAT2\*5V alleles, which were less frequent in



healthy children than in cases with leukemia. Although similar frequencies of these alleles in healthy subjects have been reported elsewhere (39, 40), we cannot discard that these results could be biased by the small sample size of the control group, which is suggested by the wide CI values observed (Table 4).

Previous studies have suggested that the rapid-acetylator genotype of NAT2 leads to an increased risk of various types of cancer, particularly leukemia, colorectal and bladder cancer

(4, 41–45). As it has been reported in Caucasian and Middle Eastern populations, rapid acetylator phenotypes were less frequent than slow acetylator phenotypes (4, 39, 46), but no statistical significance was observed between cases and controls. The predicted haplotypes described as slow acetylator TCTGAG (NAT2\*5V, CI 1.59–201.58) and rapid CTTGAG (NAT2\*11A, CI 2.05–348.8) and CTTGGG (NAT2\*12C, CI 1.32–36.83) haplotypes, which were rare in the healthy subjects (<1%),



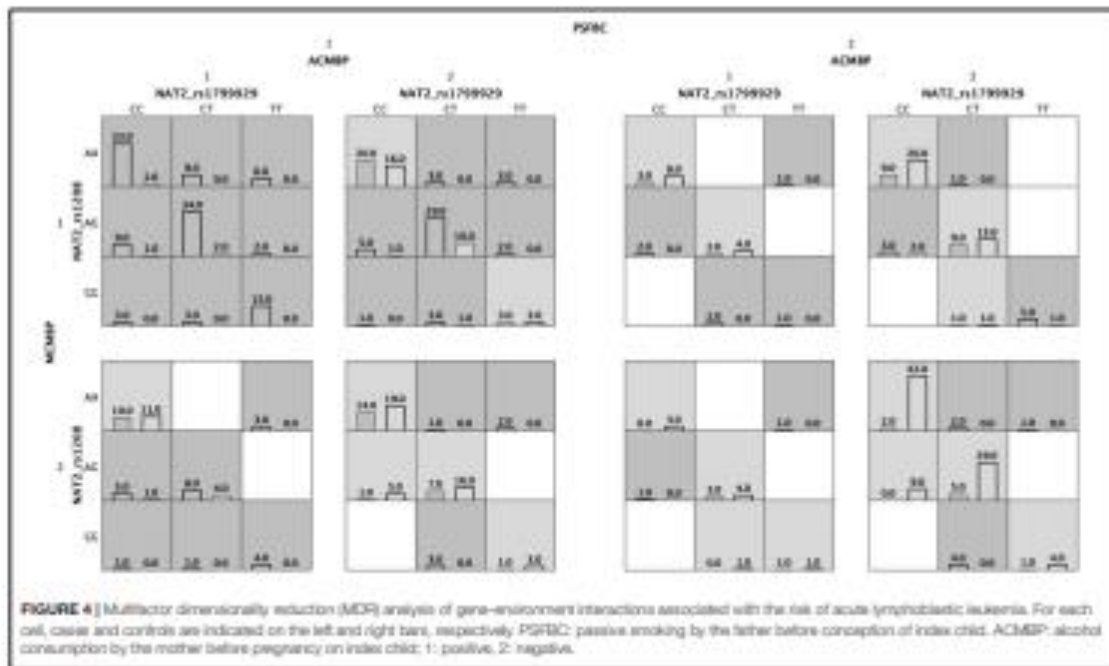
**FIGURE 3 |** Ancestry estimation based on 32 ancestry informative markers (AIMs). Principal component analysis (PCA) plot of the first two eigenvectors. Spatial ancestry distribution of acute lymphoblastic leukemia (blue), controls (pink), European (green), and Native-American (red) populations. Dispersion of samples suggests a similar genetic background between cases and controls being enriched by a Native-American contribution.

were associated with higher risk of ALL. These results seem paradoxical; however, we predicted the acetylation phenotype according to the results of the NAT2 haplotypes and potential haplotype misclassification could exist (5, 40). Uncommon haplotypes could not be clearly determined via indirect method of computational haplotype inference, however, direct methods for genotyping NAT2 haplotypes are costly and have large turnaround time. Recently, a new method has been described to eliminate potential errors in the genotypes assignment, but, this technique is based on the NAT2\*4, \*5B, \*6A, \*7B, \*12A, and \*13A, which are the six most common NAT2 haplotypes in diverse populations. Due the differences in the NAT2 SNPs and haplotype frequencies among the populations, this new method is not applicable in very heterogeneous ethnic groups, such as Mexican. In addition, it has been reported a genotype-phenotype discordance (47, 48). It is well-known that the acetylation phenotype is affected either by NAT2 genetic variants or by environmental variables.

It has been mentioned that diet and epigenetic factors could be potential modifiers of discordant association results reported by studies performed in ALL. Recent studies have suggested that dietary lifestyle has a significant influence on xenobiotic metabolism by modifying the gut microbiota and consequently NAT2 gene expression in the liver. Subjects from the same ethnic group but living in different geographic regions can be capable of responding differently to xenobiotic agents (49). Moreover, inconsistent and even contradictory data is not surprising, since multiple genes have been implicated in key pathways associated with leukemogenic processes. More over, studies of two or three genotypes in combination have also yielded inconsistent results.

It has also been hypothesized that slow acetylator phenotypes have suffered positive selection in populations under an insufficient folate diet (50). Studies in Mexican children have documented a folate deficiency in 11.2% of the children aged 4 years, which could explain our results (51). Nevertheless, our findings in the control group differ of those frequencies reported in general population from different regions of Mexico (43, 50, 52). NAT2 gene has a high frequency of functional variation that differs among ethnically diverse populations, in fact, NAT2 functional variation contributes to high levels of diversity, illustrating how geographically and temporally fluctuating xenobiotic environments may have influenced our genome variability and susceptibility to disease (52, 53). We discarded population stratification because we performed an ancestry structure analysis in a subset of patient and controls randomly selected. Our data showed that ALL cases, as well as control subjects belong to Mestizo group, the main ethnic group of Mexican population. The ancestry composition observed in the present study is in accordance with the Conquest history of Mexican population, mostly comprised by European- and Amerindian-descendent groups (54, 55).

Association analysis stratifying for exposure variables revealed that the homozygotes to the risk allele of rs1041983 (C282T) and rs1799931 (G857A) confer protection to ALL under parents exposition to diverse xenobiotics (mother: hydrocarbons, drug consumption, active smoking, alcohol consumption and insecticide; father: active smoking and alcohol drinking). To our knowledge, no previous studies have explored the relationship between NAT2 genotypes and those environmental factors in ALL. However, differential contribution of single NAT2 SNPs



to the risk of ALL has been observed recently by Zhu et al, who performed a meta-analysis including 1,522 acute leukemia patients and 2,688 controls (11).

Since ALL is considered a multifactorial disease where xenobiotics could be important factors that contribute to its pathogenesis, and that *NQO1*, *CYP2E1*, and *NAT2* are enzymes involved in the metabolism of xenobiotics (including benzene, cigarette smoking, chemotherapy agents, and alcohol drinking), which increase the risk to develop diverse human diseases (2, 11, 19), we used a MDR approach to identify combinations between genetic and environmental factors associated with ALL (20). Our data showed that the *NAT2* rs1799929-*NAT2* rs1208-*ACMBP*-*MCMBP*-*PSFBC* interact to increase the risk of ALL before pregnancy.

The *NAT2* rs1041983, *NAT2* rs1799929, *NAT2* rs1799931, *NQO1* rs1800566, *IEAB*, *ASEFAB* was the best model after child's birth, meaning that these factors in combination increase the risk of ALL (Table 5). Our data suggest that the polymorphism affect acetylation of chemicals compounds having aromatic amines as drugs, pesticides, cigarette smoke, increase the risk to develop acute ALL in Mexican children. An interaction between *NAT2* and alcohol drinking and smoking with various outcomes have been reported. It was suggested that *NAT2* could be involved in the activation of one or more pro-carcinogens associated with alcohol intake and the risk of oral squamous cell cancer (56). It is known that *NAT2* contributes to detoxification of tobacco smoke, pesticides and even prescription drugs (57–59); nonetheless, there are no evidences reporting a direct interaction within *NAT2* polymorphisms and the environmental factors identified

in the present study with the risk to develop hematological disease. The underlying mechanism for the link between genetic polymorphisms in these genes and insecticide in the development of ALL is not fully understood. A study conducted in infant leukemia with maternal exposure to dipyrone during pregnancy reported that *NAT2* SNPs are associated with this malignancy regardless of maternal exposure to the medication (39). Notwithstanding, it is well-known that studies addressing multi-gene rather than single-gene polymorphisms in xenobiotic genes could improve our knowledge of the genetic risk factors involved in ALL pathogenesis (4, 6, 25, 60).

This is a study that included patients from 8 public hospitals in Mexico City, what represents slightly <70% of all cases with leukemia in this city (61). It is also the first investigation to evaluate the interaction between *NQO1* (rs1800566), *CYP2E1* (rs3813867), and *NAT2* [rs1041983 (C282T), rs1801280 (T341C), rs1799929 (C481T), rs1799930 (G590A), rs1208 (A803G), and rs1799931 (G857A)] polymorphisms and exposure to common environmental hydrocarbons. In **Supplementary Table 4**, we can see the sociodemographic characteristics and the distribution of the hydrocarbon exposure variables of the children diagnosed with ALL of the Hospital that was not included in the present analysis. When comparing frequencies, there are no important differences to highlight, thus reducing the possibility of selection bias.

On the other hand, with respect to the variables included in this study to assess the exposure to hydrocarbons, we used a similar strategy reproduced in previous studies where the association between the exposure to hydrocarbons and the risk

**TABLE 5 |** Gene-environment interaction and acute lymphoblastic leukemia risk prediction.

Exposure Window	# Factors	Model	TA	TBA	CVC	OR (CI 95%)	P-value*
Before pregnancy	1	PSFDC	0.7087	0.7087	10/10	8.48 (1.90-22.08)	<0.0001
	2	MCMER, PSFDC	0.7245	0.6952	8/10	8.19 (5.28-12.52)	<0.0001
	3	NAT2_rs1799209, MCMER, PSFDC	0.7460	0.7081	5/10	8.58 (5.78-12.74)	<0.0001
	4	NAT2_rs1799209, NAT2_rs1208, MCMER, PSFDC	0.7773	0.7329	8/10	11.70 (7.72-17.72)	<0.0001
	5	NAT2_rs1799209, NAT2_rs1208, ACMER, MCMER, PSFDC	0.7905	0.7648	10/10	18.98 (11.72-30.74)	<0.0001
	6	NAT2_rs1799209, NQO1_rs1800595, CYP2E1_rs1803867, ACMER, MCMER, PSFDC	0.8248	0.8036	3/10	19.04 (12.28-29.52)	<0.0001
	7	NAT2_rs1041983, NAT2_rs1208, NQO1_rs1800595, MD4ER, ACMER, MCMER, PSFDC	0.8588	0.8145	5/10	29.31 (21.74-40.12)	<0.0001
	8	NAT2_rs1041983, NAT2_rs1208, NQO1_rs1800595, MD4ER, ACMER, ACMER, MCMER, ASFAB, PSFDC	0.8933	0.8382	2/10	75.72 (60.38-122.88)	<0.0001
During Pregnancy	1	PSFDP	0.6889	0.6889	10/10	8.88 (4.43-10.07)	<0.0001
	2	ASMR, PSFDP	0.6987	0.6794	7/10	7.31 (4.94-11.42)	<0.0001
	3	NAT2_rs1799209, ASMR, PSFDP	0.7141	0.6820	5/10	8.05 (5.34-12.18)	<0.0001
	4	NAT2_rs1041983, NAT2_rs1799209, NQO1_rs1800595, PSFDP	0.7318	0.6491	4/10	7.15 (4.95-10.45)	<0.0001
	5	NAT2_rs1041983, NAT2_rs1799209, NQO1_rs1800595, CYP2E1_rs1803867, PSFDP	0.7932	0.6252	6/10	12.19 (8.82-15.23)	<0.0001
	6	NAT2_rs1041983, NAT2_rs1799209, NAT2_rs1799201, NQO1_rs1041983, CYP2E1_rs1803867, PSFDP	0.7945	0.6296	7/10	14.30 (9.90-22.63)	<0.0001
	7	NAT2_rs1041983, NAT2_rs1799209, NAT2_rs1799201, NQO1_rs1041983, CYP2E1_rs1803867, HDNDR, PSFDP	0.8258	0.6196	5/10	22.93 (14.88-35.90)	<0.0001
	8	NAT2_rs1041983, NAT2_rs1799209, NAT2_rs1799201, NQO1_rs1041983, CYP2E1_rs1803867, HDNDR, ASMR, PSFDP	0.8608	0.6774	4/10	25.79 (22.15-27.88)	<0.0001
After birth	1	NAT2_rs1799209	0.6021	0.6021	10/10	2.29 (1.62-3.22)	<0.0001
	2	NAT2_rs1799209, ASMR	0.6296	0.6095	8/10	3.21 (2.08-4.18)	<0.0001
	3	NAT2_rs1041983, NAT2_rs1799209, NAT2_rs1799201	0.6561	0.606	8/10	3.52 (2.48-5.01)	<0.0001
	4	NAT2_rs1041983, NAT2_rs1799209, NAT2_rs1799201, NQO1_rs1800595	0.6848	0.5889	5/10	4.54 (3.17-6.51)	<0.0001
	5	NAT2_rs1041983, NAT2_rs1799209, NAT2_rs1799201, NQO1_rs1800595, IEAB	0.7109	0.5724	5/10	3.88 (2.92-4.17)	<0.0001
	6	NAT2_rs1041983, NAT2_rs1799209, NAT2_rs1799201, NQO1_rs1800595, IEAB, ASFAB	0.7553	0.6082	8/10	9.22 (6.24-13.61)	<0.0001

Best models gene-environment interactions with cross-validation consistency (CVC) and prediction error per individual model that have been obtained by the multivariate dimensionality reduction (MDR) of our data set. Table is displaying best models with TBA > 0.57. \*Values from Whole-Site for data set. TA, training balance accuracy; TBA, testing balance accuracy; PSFDC, passive smoking by the father before conception; MCMER, Medicine consumption by the mother before pregnancy; ACMER, alcohol consumption by the mother before pregnancy; MD4ER, hydrocarbons exposure by the mother before pregnancy; ASMR, active smoking by the mother during pregnancy; ASMRP, active smoking by the mother before pregnancy; PSFDP, passive smoking by the father during pregnancy; ASFP, active smoking by the father during pregnancy; HDNDR, hydrocarbon exposure by the mother during pregnancy; ASMRB, active smoking by the mother after child's birth; IEAB, insecticide exposure of the mother after birth; ASFAB, active smoking by the father after birth.

of developing childhood leukemia has been analyzed (62-65). This study had the disadvantage that it was limited to what the cases and controls remember, but our instrument included a question if the parents of the cases identified whether a possible cause of the disease was exposure to hydrocarbons, tobacco or alcohol consumption. There were no responses related to this. Thus, it can be noted that there is little likelihood of recall bias associated with the measurement of these variables. Another point in favor is that neither the interviewers nor the interviewees knew the results of the polymorphisms, so finding the existence of interactions between some variables and the polymorphisms reinforces the strategy with which our variables were measured. A logistic regression model was performed to evaluate the variables that could potentially be confusing, which is shown in Table 1. This was complemented with a recessive analysis which allowed us to adjust for possible confounding variables.

There were some limitations in the present study. On one hand, we tested only one SNP of *NQO1* and *CYP2E1* genes, thus we cannot discard the association among these genes and ALL or the interaction among SNPs in these genes and environmental xenobiotics. On the other hand, subjects were grouped into three different *NAT2* acetylator phenotypes based only on six *NAT2* SNPs (slow acetylators: two slow alleles, intermediate acetylators: one slow and one rapid allele, and rapid acetylators: 2 rapid alleles). Considering that we did not study the rs1801279 (191G>A), the assessment of only these six SNPs could result in a misclassification of some *NAT2* alleles (39). In addition, new alleles of *NAT2* have been identified and there is there is phenotype heterogeneity within the slow and intermediate acetylator genotype groups, due to variation in enzyme activity conferred by different alleles (66). Low genotype rate could bias our phenotype results; nevertheless, it has been reported

that using only the rs1041983 (C282T), rs1801280 (T341C) SNPs, it is possible to predict the NAT2 phenotype with high sensitivity and specificity (0.9993 and 0.9880, respectively) in Caucasian, Latin-American and Middle East populations (42, 67). Otherwise, controversial results in NAT2 association findings among studies could be explained by factors, such as sample size, age group, genotyping method, and the time of exposure to risk agents (7, 68, 69).

## CONCLUSION

To the best of our knowledge, this is the first assessment of the interaction between hydrocarbon exposure and genetic polymorphisms of NAT2, NQO1, and CYP2E1 on the risk of childhood ALL in the Mexican population, and the first report of ancestry background in Mexican children with ALL. Our study provides evidence that polymorphisms of NAT2 might be genetic factors involved in childhood ALL. These results shed light on the contribution of NAT2 polymorphisms to increase the risk of developing ALL in children.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics and National Committee of Scientific Research of the Instituto Mexicano del Seguro Social approved this study with number R-2013-785-062. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

AM-S, JM-A, and SJ-M: conceptualization. AM-S, JN-E, JCF-L, JM-A, and SJ-M: methodology. JN-E, JCF-L, JM-A, and SJ-M: formal analysis. AM-S, JN-E, and SJ-M: investigation. AM-S, EH-C, MP-S, AM-G, EJ-H, JM-T, HP-L, JF-L, RA-S, FM-R, JGP-G, DD-R, JT-N, JF-B, RE-E, PR-Z, LF-V, JG-U, SM-S, GE-A, CA-H, RR-C, LH-M, LG-L, GC-O, AG-E, IC-H, AM-H, ML-C,

NH-P, JG-K, MR-V, DT-V, CC-R, FM-L, JAP-G, AM-R, AA-D, BS-D, VB-M, MM-R, OS-R, JR-R, HR-V, and AH-M: resources. AM-S and SJ-M: writing—original draft preparation. SJ-M and JM-A: writing—review and editing, supervision, and funding acquisition. All authors reviewed the final manuscript, read, and approved the submitted version.

## FUNDING

This work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT), Investigación en Fronteras de la Ciencia (IFC)-2016-01-2119, PDCPN2013-01-215726, FIS/IMSS/PROT/1364; SALUD 2015-1-262190, FIS/IMSS/PROT/1533; CB-2015-1-258042, FIS/IMSS/PROT/1548, and FONCICYT/37/2018, FIS/IMSS/PROT/1782. We also thank for the financial support from the National Institute of Genomic Medicine (19/2019/I) and the Instituto Mexicano del Seguro Social, grant numbers: FIS/IMSS/PROT/PRIO/14/031, FIS/IMSS/PROT/PRIO/15/048, FIS/IMSS/PROT/PRIO/18/080, and FIS/IMSS/PROT/PRIO/19/088. We also thank for the financial support from the National Institute of Genomic Medicine (19/2019/I) and the Hospital Infantil de México Federico Gómez through the Fondo de apoyo a la Enseñanza e Investigación en Oncología Pediátrica and the Programa de Fondos Federales para la Investigación del Hospital Infantil de México Federico Gómez.

## ACKNOWLEDGMENTS

We were very grateful with Dr. Catherine Metayer and MPH Alice Y. Kang, University of California, Berkeley School of Public Health, USA, because of their donation of saliva kits to perform this project. As well, we thank to María del Carmen Alvarez-Hernández, hospital staff and all the patients and their families whose willingness to participate made this study possible to be conducted and successfully finished. This manuscript is the product of the Ph.D. thesis of AM-S within the Programa de Maestría y Doctorado en Ciencias Médicas de la Facultad de Medicina, UNAM.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.571869/full#supplementary-material>

## REFERENCES

- Mejía Anagari JM, Ortega Alvarez MC, Ejarque Gutiérrez A. Acute leukemias epidemiology in children. Part 1. *Rev Med Inst Mex Seguro Soc.* (2007) 43:323–33.
- Krajčović M, Labuda D, Mathonnet G, Labuda M, Moghribi A, Champagne J, et al. Polymorphisms in genes encoding drug and xenobiotic metabolizing enzymes, DNA repair enzymes, and response to treatment of childhood acute lymphoblastic leukemia. *Clin Cancer Res.* (2002) 8:882–93.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* (2000) 157:945–59.
- Krajčović M, Richer C, Sinnett H, Labuda D, Sinnett D. Genetic polymorphisms of N-acetyltransferase 1 and 2 and gene-gene interaction

- in the susceptibility to childhood acute lymphoblastic leukemia. *Cancer Epidemiol Biomarkers Prev.* (2000) 9:527–42.
- Hain DW, Doll MA, Forland AJ, Loff MA, Webb SJ, Xiao GH, et al. Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. *Cancer Epidemiol Biomarkers Prev.* (2000) 9:29–42.
- Basma HA, Kobissi LH, Jabbar ME, Mousa MA, Dhaini HR. CYP2E1 and NQO1 genotypes and bladder cancer risk in a Lebanese population. *Int J Mol Epidemiol Genet.* (2013) 4:207–17.
- Jiang WH, Wang XT, Zhang LD, Yan QJ, Chen LL. Relationship between NAT2 polymorphisms and onset risk of acute leukemia: a systematic review and meta-analysis. *Eur Rev Med Pharmacol Sci.* (2019) 23:9109–46. doi: 10.26355/eurresv\_201901\_0408



47. Kotila OA, Feroz OI, Olajade OI, Ayode AI, Fajana AG, Babalola CP. N-acetyltransferase 2 enzyme genotype-phenotype discordance in both HIV-negative and HIV-positive Nigerians. *Pharmacogenomics Genomics*. (2019) 20:186–93. doi: 10.1097/FPG.0000000000000073
48. O'Shea WM, Doublich RK, MacArthur RD, Fureough MJ, Doll MA, Fritland AJ, et al. Acetylator phenotype and genotype in patients infected with HIV. Discordance between methods for phenotype determination and genotype. *Pharmacogenomics*. (2000) 10:171–82. doi: 10.1097/00008771-200003000-00009
49. Djordjevic N, Carrillo JA, Roh HK, Karlsson S, Ueda N, Bertilsson L, et al. Comparison of N-acetyltransferase 2 enzyme genotype-phenotype and xanthine oxidase enzyme activity between Sweden and Kansas. *J Clin Pharmacol*. (2012) 52:1327–34. doi: 10.1177/0891220111420261
50. Valente C, Alvarez L, Marks SJ, Lopez-Perez AM, Paraiso W, Ouchafout O, et al. Exploring the relationship between HbA1c, beta and genetic adaptations in humans. *BMC Genet*. (2015) 16:55. doi: 10.1186/s12863-015-0212-1
51. Cuevas-Nasu L, Mendoza-Soto Y, Shamsh-Lavy T, Mendoza-Gilman-Hernandez I, Ariza-Arreola MA, Robledo-Campos Mdel R, et al. Prevalence of beta and alpha R12 deficiency in Mexican children aged 1 to 6 years in a population-based survey. *Salud Publica Mex*. (2012) 54:116–24. doi: 10.1590/S0034-36342012000200007
52. Tajo-Chayeh L, Aguirre JA, Miguero-Martin C, Chavez-Blanco A, Duarte-Gonzalez A. Arylamine N-acetyltransferase 2 genotypes in a Mexican population. *Gene Mol Res*. (2012) 11:382–92. doi: 10.4230/2012.April.27.7
53. Awanthakrishnan AN, Du M, Baruch SI, Brenner H, Cain RL, Casey G, et al. Red meat intake, NAT2, and risk of colorectal cancer: a pooled analysis of 13 studies. *Cancer Epidemiol Biomarkers Prev*. (2015) 24:598–605. doi: 10.1158/1075-9965.EPI-14-0887
54. Jimenez-Monreal S, Ramirez-Flanencia M, Mejia-Aranguez JM, Nieto-Escribano JC, Bekker-Mendez C, Torres-Escalante R, et al. Analysis of thiopurine S-methyltransferase deficient alleles in acute lymphoblastic leukemia patients in Mexican patients. *Archiv Med Res*. (2006) 47:615–22. doi: 10.1016/j.arcmed.2016.11.018
55. Paris E, Barreiro LB, Saberi PC, Anagnostis E, Laca E, Szejnfar A, et al. Deciphering the ancient and complex evolutionary history of human arylamine N-acetyltransferase genes. *Am J Hum Genet*. (2006) 78:623–36. doi: 10.1086/508614
56. Chen C, Ricci S, Doody DR, Fitzgibbon ED, Porter PL, Schwartz SM. N-Acetyltransferase 2 polymorphisms, cigarette smoking and alcohol consumption, and oral squamous cell cancer risk. *Carcinogenesis*. (2001) 22:1993–9. doi: 10.1093/carcin/22.12.1993
57. Taylor MC, Small CM, Dominguez CE, Murray LR, Tang W, Wilson MM, et al. Alcohol, smoking, and caffeine in relation to fecundability, with effect modification by NAT2. *Am J Epidemiol*. (2011) 21:866–72. doi: 10.1093/aje/kwq181
58. Weber WW, Hein DW. N-acetylation pharmacogenetics. *Pharmacol Rev*. (1985) 37:25–79.
59. Zambrano CM, Escarotichan M, Gonzalez BA, Faco A, Kothman S, Pombo-de-Oliveira MS. N-acetyltransferase 2 polymorphisms and susceptibility to infant leukemia with maternal exposure to dipyridone during pregnancy. *Cancer Epidemiol Biomarkers Prev*. (2008) 19:3037–41. doi: 10.1158/1075-9965.EPI-08-0508
60. Kuznetsov M, Simon H, Richer C, Labuda D, Simon D. Role of NQO1, MPO and CYP2E1 genetic polymorphisms in the susceptibility to childhood acute lymphoblastic leukemia. *Int J Cancer*. (2002) 97:230–6. doi: 10.1002/ijc.1589
61. Martin-Trejo JA, Nieto-Escribano JC, Fajardo-Gutierrez A, Medina-Sanson A, Flores-Lajons J, Jimenez-Hernandez E, et al. Early mortality in children with acute lymphoblastic leukemia in a developing country: the role of malnutrition at diagnosis. A multicenter cohort MIGHOCL study. *Leuk Lymphoma*. (2017) 58:998–998. doi: 10.1080/10428194.2016.1211904
62. Bakky HD, Infante-Rivero C, Menayor C, Clavel J, Lightfoot T, Katusch R, et al. Home pesticide exposures and risk of childhood leukemia: Findings from the childhood leukemia international consortium. *Int J Cancer*. (2015) 137:2644–51. doi: 10.1002/ijc.29631
63. Menayor C, Petrakou E, Aranguez JIM, Roman E, Schüz J, Magrassi C, et al. Parental tobacco smoking and acute myeloid leukemia: The Childhood Leukemia International Consortium. *Am J Epidemiol*. (2004) 160:261–71. doi: 10.1093/aje/kwh028
64. Perez-Saldívar ML, Ortega-Alvarez MC, Fajardo-Gutierrez A, Benabib-Escobedo R, Del Campo-Martinez Mde L, Medina-Sanson A, et al. Father's occupational exposure to carcinogenic agents and childhood acute leukemia: a new method to assess exposure (a case-control study). *BMC Cancer*. (2008) 8:7. doi: 10.1186/1471-2407-8-7
65. Perez-Saldívar ML, Fajardo-Gutierrez A, Sierra-Bautista JA, Nieto-Villagas N, Feroz-Lomaxana H, Domínguez-Acosta EM, et al. Parental exposure to workplace carcinogens and the risk of development of acute leukemia in infants. Case-control study. *Archiv Med Res*. (2006) 47:684–93. doi: 10.1016/j.arcmed.2016.12.006
66. McDonough EM, Boskovic S, Akiba E, Heis DW, Altman RB, Klein TE. PharmGKB summary: very important pharmacogenetics information for N-acetyltransferase 2. *Pharmacogenomics Genomics*. (2014) 15:489–25. doi: 10.1097/FPG.0000000000000082
67. Sefidki S, Hlaszkovics M, Lefrançois ML, Ovtziankhar D, Moermann C, Guballa C, et al. Genotyping NAT2 with only two SNPs (rs1841983 and rs1801280) outperforms the tagging SNP rs1895741 and is equivalent to the conventional 7-SNP NAT2 genotype. *Pharmacogenomics Genomics*. (2011) 12:673–4. doi: 10.1097/FPG.0b013e3182149323
68. Gökçe K, Hlaszkovics M, Sefidki S, Hengstler JG, Bek HM. NAT2 and Madras cancer letter. *Cancer Epidemiol Biomarkers Prev*. (2004) 13:561. doi: 10.1158/1075-9965.EPI-03-1261
69. Pesch B, Gaszych K, Rabstein S, Weiss T, Casjes S, Riba HP, et al. N-acetyltransferase 2 phenotype, occupation, and bladder cancer risk: results from the EPIC cohort. *Cancer Epidemiol Biomarkers Prev*. (2013) 22:2055–65. doi: 10.1158/1075-9965.EPI-13-0119-7

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Citation: Medina-Sanson A, Nieto-Escribano JC, Hurtado-Cordero E, Pérez-Saldívar ML, Martínez-García A, Jiménez-Hernández E, Fernández-López JC, Martín-Trejo JA, Feroz-Lomaxana H, Flores-Lajons J, Amador-Sánchez R, Mora-Ríos PG, Pedraza-González JI, Duarte-Rodríguez DA, Torres-Nieto JR, Flores-Bautista JA, Espinosa-Estrada RM, Acosta-Zepeda JM, Flores-Villagas JI, González-Ulvérrri JE, Martínez-Silva R, Espinosa-Arreola C, Alameda-Hernández C, Ramírez-Colonado R, Hernández-Mora L, García-López LR, Cruz-Oyola GA, Gadey-Español AE, Contreras-Hernández J, Medina-Hernández A, López-Caballero MG, Hernández-Pineda NA, González-Knaulff J, Rodríguez-Vilpaz MA, Torres-Velazco D, Cortés-Rojas C, Medina-López FJ, Pérez-Gómez JA, Martínez-Ríos A, Aguilar-De la Serna A, Scoglio-Díaz B, Bekker-Méndez FC, Mata-Rocha M, Monzó-Castillo BA, Sepúlveda-Robles DA, Ramírez-Bello J, Rosas-Vargas H, Hidalgo-Miranda A, Mejía-Aranguez JM and Jiménez-Monreal S (2020) Genotype-Environment Interaction Analysis of NQO1, CYP2E1, and NAT2 Polymorphisms and the Risk of Childhood Acute Lymphoblastic Leukemia: A Report From the Mexican Interenvironmental Group for the Identification of the Causes of Childhood Leukemia. *Front Oncol* 10:571968. doi: 10.3389/fonc.2020.571968

Copyright © 2020 Medina-Sanson, Nieto-Escribano, Hurtado-Cordero, Pérez-Saldívar, Martínez-García, Jiménez-Hernández, Fernández-López, Martín-Trejo, Feroz-Lomaxana, Flores-Lajons, Amador-Sánchez, Mora-Ríos, Pedraza-González, Duarte-Rodríguez, Torres-Nieto, Flores-Bautista, Espinosa-Estrada, Rosas-Zepeda, Flores-Villagas, González-Ulvérrri, Martínez-Silva, Espinosa-Arreola, Alameda-Hernández, Ramírez-Colonado, Hernández-Mora, García-López, Cruz-Oyola, Gadey-Español, Contreras-Hernández, Medina-Hernández, López-Caballero, Hernández-Pineda, González-Knaulff, Rodríguez-Vilpaz, Torres-Velazco, Cortés-Rojas, Medina-López, Pérez-Gómez, Martínez-Ríos, Aguilar-De la Serna, Scoglio-Díaz, Bekker-Méndez, Mata-Rocha, Monzó-Castillo, Sepúlveda-Robles, Ramírez-Bello, Rosas-Vargas, Hidalgo-Miranda, Mejía-Aranguez and Jiménez-Monreal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.