Annals of Hematology (2019) 98:2379–2388
https://doi.org/10.1007/s00277-019-03730-x

ORIGINAL ARTICLE

Variants in *ARID5B* gene are associated with the development of acute lymphoblastic leukemia in Mexican children

Adriana Reyes-León[1] · Maribel Ramírez-Martínez[1] · Diana Fernández-García[1] · David Amaro-Muñoz[1] · José Antonio Velázquez-Aragón[2] · Consuelo Salas-Labadía[1] · Marta Zapata-Tarrés[3] · Liliana Velasco-Hidalgo[3] · Norma López-Santiago[4] · Mayra Ivette López-Ruiz[5] · Monica Anabell Malavar-Guadarrama[6] · Rocío Cárdenas-Cardós[3] · Rogelio Paredes-Aguilera[4] · Roberto Rivera-Luna[7] · Michael Dean[8] · Patricia Pérez-Vera[1]

Received: 3 January 2019 / Accepted: 10 June 2019 / Published online: 21 June 2019

Abstract

A high impact of *ARID5B* SNPs on acute lymphoblastic leukemia (ALL) susceptibility has been described in Hispanic children; therefore, it is relevant to know if they influence the high incidence of childhood-ALL in Mexicans. Seven SNPs (rs10821936, rs10994982, rs7089424, rs2393732, rs2393782, rs2893881, rs4948488) of *ARID5B* were analyzed in 384 controls and 298 ALL children using genomic DNA and TaqMan probes. The SNPs were analyzed for deviation of Hardy-Weinberg equilibrium; Fisher's exact test was used to compare the genotypic and allelic frequencies between controls and patients. The association between SNPs and ALL susceptibility was calculated, and haplotype and ancestry analyses were conducted. All SNPs were associated with ALL, pre-B ALL, and hyperdiploid-ALL susceptibility ($p < 0.05$). No association with T-ALL and gene fusions was found ($p > 0.05$). The seven SNPs were associated with risk of pre-B ALL in younger children; however, rs2393732, rs2393782, rs2893881, and rs4948488 were not associated with susceptibility in older children and adolescents. The CAG haplotype (rs10821936, rs10994982, rs7089424) was strongly associated with ALL risk in our population ($p < 0.00001$). The frequency of all risk alleles in our ALL, pre-B, and hyperdiploid-ALL patients was higher than that in Hispanic children reported. This is the first report showing the association between rs2393732, rs2393782, and rs4948488 with pre-B hyperdiploid-ALL children. The G allele at rs2893881 confers major risk for pre-B hyperdiploid-ALL in Mexican (OR, 2.29) than in Hispanic children (OR, 1.71). The genetic background of our population could influence the susceptibility to ALL and explain its high incidence in Mexico.

Keywords Acute lymphoblastic leukemia · ARID5B gene · Single nucleotide polymorphisms · Susceptibility · Mexican children

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00277-019-03730-x) contains supplementary material, which is available to authorized users.

✉ Patricia Pérez-Vera
pperezvera@yahoo.com

[1] Laboratorio de Genética y Cáncer, Instituto Nacional de Pediatría, Insurgentes Sur 3700-C, Colonia Insurgentes Cuicuilco, C.P. 04530 Mexico City, Mexico

[2] Laboratorio de Biología Molecular, Instituto Nacional de Pediatría, Mexico City, Mexico

[3] Servicio de Oncología, Instituto Nacional de Pediatría, Mexico City, Mexico

[4] Servicio de Hematología, Instituto Nacional de Pediatría, Mexico City, Mexico

[5] Departamento de Oncología, Hospital de Especialidades Pediátricas de Tuxtla Gutiérrez, Tuxtla Gutiérrez, Chiapas,, Mexico

[6] Departamento de Hematología, Hospital de Especialidades Pediátricas de Tuxtla Gutiérrez, Tuxtla Gutiérrez, Chiapas, Mexico

[7] Subdirección de Hemato-Oncología, Instituto Nacional de Pediatría, Mexico City, Mexico

[8] Laboratory of Translational Genomics, Division of Cancer Epidemiolgy & Genetics, National Cancer Institute, Bethesda, MD, USA

Introduction

Acute lymphoblastic leukemia (ALL) is characterized by the uncontrolled proliferation and malignant transformation of lymphoid progenitor cells in the bone marrow. It is the most common subtype of leukemia diagnosed in children under 18 years [24].

The incidence of childhood ALL in the United States of America (USA) varies according to race or ethnic group. For example in Hispanic children, ALL incidence is higher (40.9/1,000,000/year) than in Asian (35.6/1,000,000/year), White (33/1,000,000/year), and Black (16/1,000,000/year) children [23, 41]. In Mexico, cancer is the second leading cause of childhood mortality and ALL represents 83% of the leukemia cases. According to the statistics of the Popular Medical Insurance Program (PMIP), the national ALL incidence is 79.8 cases per million per year and it is significantly increasing [32].

ALL is a multifactorial disease, in which environmental (prenatal exposure to X-rays, pesticides, and radiation), biological (viral infections), and genetic (Bloom syndrome, ataxia telangiectasea, neurofibromatosis, B-thalassemia, and constitutional trisomy 21) factors play an important role in the development of the disease [7, 10]. However, these factors are not enough to explain the etiology of ALL and the high incidence of childhood ALL in our country. In addition to these factors, genetic variations such as single nucleotide polymorphisms (SNPs) could also contribute to the development of the disease in children. Recently, the presence of several SNPs in genes including *IKZF1* (7p12.2), *ARID5B* (10q21.2), *DDC* (7p12.1), *CEBPE* (14q11.2), *CDKN2A* (9p21.3), *GATA3* (10p14), and *PIP4K2A* (10p12.2) have been associated with the susceptibility to develop ALL in different populations around the world [5, 13, 14, 20, 25–27, 30, 36, 38, 40, 41]. Apparently, the treatment outcome, prognosis, and overall survival of childhood ALL depend on race (Asian, Black, White, and Native American) or ethnicity (Hispanic, and Non-Hispanic) [6, 18].

Interestingly, specific SNPs of *ARID5B* (rs7089424, rs10994982, rs10740055, rs10821936, rs2393732, rs17215180, and rs2393782), *CDKN2A* (rs3218018, rs3731249, and rs3731217), *IKZF1* (rs4132601 and rs7780012), *CEBPE* (rs4982731, rs10143875, rs2239633, and rs2239635), and *PIP4K2A* (rs7901152, rs11013046, rs7088318, and rs7075634) are more associated with ALL risk in Hispanic than in White children [8, 16, 37, 39, 40]. Furthermore, some of them (rs10821936, rs7089424, rs7088318, rs2239633, and rs3731217) are highly correlated with the increase of Native American ancestry [37, 39]. Although the C risk allele of rs10821936 has been associated with susceptibility to childhood ALL in both Hispanic and White populations, the allelic frequency was higher in Hispanic (0.47) than in White (0.33) and Black (0.16) children, showing that there are differences among populations [39].

Although some SNPs of *ARID5B* have been more related with ALL risk in Hispanic children, these observations are not necessarily representative of the situation in Mexico, since the proportion of Mexican patients included in these studies is unknown. It is essential for us to know if the genetic variants of *ARID5B* confer susceptibility to the development of the disease in our population and if these variants contribute to the higher incidence of childhood ALL in Mexico. This study helps to understand the biologic mechanisms involved in the etiology of childhood ALL considering the Mexican genetic background. The aim of this study was to determine the association between the presence of the risk alleles of the seven SNPs of *ARID5B* gene and the susceptibility to develop ALL in Mexican children.

Materials and methods

Study population

A total of 298 patients under 18 years of age with a diagnosis of ALL (precursor B cells or T cells) were studied (Table 1). Patients were recruited at any stage of the disease (diagnosis, treatment, relapse, and surveillance) between August 2012 to July 2016 at the Servicio de Hematología y Oncología del Instituto Nacional de Pediatría in Mexico City, and the Hospital de Especialidades Pediátricas de Tuxtla Gutiérrez, Chiapas, Mexico. The diagnosis was established by cytomorphology, immunophenotyping, molecular biology for the most common translocations, and DNA index. The blast lineage was determined by the expression of different CD (cluster of differentiation) cell surface markers; the expression of CD10, CD19, CD20, and CD22 defines pre-B lineage, and CD1, CD2, CD3, CD5, and CD7 defines T-lineage. The patient's clinical and laboratory data were obtained from the clinical records.

The control group consisted of 384 randomly selected, healthy, unrelated volunteers with no family history of hematological cancer. They were recruited from August 2013 to July 2016. Personal information was obtained through a direct interview with the volunteers, or with parents or legal tutors for individuals younger than 18 years old.

Patients and controls were Mexican mestizo residents in our country, with parents and grandparents born in Mexico.

This study was reviewed and approved by the Institutional Research and Ethics Committees from both participant Institutions in accordance with the ethical principles of the Declaration of Helsinki. Volunteers, parents, or legal tutors were previously informed about the study, and before biological samples were collected, they provided a signed, written informed consent letter to participate.

Table 1 Clinical characteristics of the patients

Clinical characteristics	pre-B ALL (284 patients) *n* (%)	T-ALL (14 patients) *n* (%)	Total (298 patients) *n* (%)
Gender			
Male	166 (58.5)	12 (85.7)	178 (59.7)
Female	118 (41.5)	2 (14.3)	120 (40.3)
Age at diagnosis (years)			
1–9	188 (66.2)	7 (50)	195 (65.5)
< 1 or ≥ 10	96 (33.8)	7 (50)	103 (34.5)
White blood cell (WBC) count			
$< 50 \times 10^9$/L	228 (80.3)	3 (21.4)	231 (77.5)
$\geq 50 \times 10^9$/L	56 (19.7)	11 (78.6)	67 (22.5)
Gene fusions			
Negative	246 (86.6)	12 (85.8)	258 (86.7)
Positive	38 (13.4)	2 (14.2)	40 (13.3)
ETV6-RUNX1	15 (5.3)	0	15 (5)
TCF3-PBX1	12 (4.2)	0	12 (4)
BCR-ABL1	6 (2.1)	0	6 (2)
KMT2A-MLLT3	2 (0.7)	0	2 (0.7)
KMT2A-EPS15	2 (0.7)	0	2 (0.7)
KMT2A-AFF1	1 (0.4)	0	1 (0.3)
TAL-1	0	1 (7.1)	1 (0.3)
TRD-LMO2	0	1 (7.1)	1 (0.3)
Numerical abnormalities			
Diploid	183 (64.4)	6 (42.8)	189 (63.4)
Hypodiploid	22 (7.7)	6 (42.8)	28 (9.4)
Hyperdiploid	39 (13.7)	0	39 (13.1)
Unknown	40 (14.1)	2 (14.3)	42 (14.1)

ARID5B genotyping analysis

Saliva samples were collected from all patients according to commercial kit instructions (Oragen DNA kit, DNA Genotek Inc. Ottawa, ON, Canada). Peripheral blood samples from controls were collected in EDTA-supplemented tubes. Saliva and peripheral blood samples were processed according to the genomic DNA extraction protocols prepIT-L2P kit (DNA Genotek Inc. Ottawa, ON, Canada) and QIAamp DNA Blood kit (QIAGEN, Hilden, Germany), respectively. Genomic DNA integrity was confirmed by 1% agarose gel electrophoresis stained with GelRed (Nucleic Acid Gel Stain, Biotium Inc. Fremont, CA, USA) and visualized in a transilluminator (GelDoc-It Imaging Systems, UVP, LLC. CA, USA). Genomic DNA was quantified using an Epoch Microplate Spectrophotometer (BioTek Instruments, Inc. Winooski, VT, USA).

Genotyping analysis was performed by real-time polymerase chain reaction (RT-PCR) in a StepOne Real-Time PCR equipment (Applied Biosystems, Foster City, CA, USA) under standard conditions. Seven SNPs (rs10821936, rs10994982, rs7089424, rs2393732, rs2393782, rs2893881, and rs4948488) of *ARID5B* gene were genotyped with a pre-designed TaqMan assay (VIC/FAM dye-labeled fluorescent probes; Applied Biosystems, Foster City, CA, USA). Total volume of the amplification reaction was 25 μL, containing 20 ng of genomic DNA, 2x TaqMan Universal PCR Master Mix No AmpErase UNG (Applied Biosystems, Foster City, CA, USA), 20x TaqMan genotyping assay-human probe (Applied Biosystems, Foster City, CA, USA), and DNAse-free sterile-filtered water. The thermal cycle program consisted of an initial denaturation at 60 °C for 30 s and 95 °C for 10 min, followed by amplification step of 40 cycles of 15 s at 95 °C and 60 s at 60 °C; final step consisted of 30 s at 60 °C. Each experiment included one negative control and three positive controls (allele 1 homozygous, heterozygous, and allele 2 homozygous). Amplification was repeated on 30% of the samples randomly, and absolute replication and concordance were observed.

Ancestry analysis

Ten ancestry informative markers (AIMs) rs2695, rs2862, rs3340, rs17203, rs203096, rs223830, rs722098, rs1800498, rs2065160, and rs2814778 were used for

ancestry analysis in patients (KASP assay); this panel has been previously used in case-control studies to distinguish the European and Native American ancestry ($d > 0.44$) in Mexican population [22]. For the control group, the ancestry analysis was carried out with the genotypes obtained from the Infinium OncoArray-500K (Illumina, Inc. USA).

To estimate the proportion of European and Native American ancestry in each patient and control, the STRUCTURE program v2.3 was utilized [17]. The parameters of the analysis were as follows: a burn-in period of 500,000 repetitions, followed by an analysis period of 500,000 repetitions under the admixed model with a value of $K = 2$ (K = ancestral populations). The GLU struct.admix module (http://code.google.com/p/glu-genetics/) was used to define the ancestry proportions among controls. The ancestry proportions of the controls and patients were compared using the Student t test, and p values less than 0.05 were considered statistically significant.

Statistical analysis

All SNPs were analyzed for deviation from Hardy-Weinberg equilibrium (HWE) (DeFinetti software; https://ihg.gsf.de/cgi-bin/hw/hwa2.pl), and the genotype and allelic frequencies were calculated. Two-tailed Fisher's exact test (GraphPad Software, Inc. La Jolla, CA, USA) was used to calculate the genotype distribution in controls and patients, and to compare the differences between both groups; p values less than 0.05 were considered statistically significant. The association of each SNP with the susceptibility to ALL, pre-B-hyperdiploid ALL, and gene fusions was determined by the p value. Odds ratio (OR) with 95% confidence intervals (95% IC) (DeFinetti software; https://ihg.gsf.de/cgi-bin/hw/hwa2.pl) was calculated to estimate the risk of developing childhood ALL in the presence of risk alleles.

Seven SNPs were submitted to haplotype analysis using the Haploview program [3] to know if they were in linkage disequilibrium (LD) ($D' > 0.80$); the haplotype blocks formed were submitted to association analysis with susceptibility of ALL; the p values were obtained by permutation analysis (100,000 repetitions). A positive association with ALL susceptibility was considered with a p value < 0.05.

Data availability The datasets generated during the current study are available from the corresponding author on reasonable request.

Results

Characteristics of controls and patients

Of the 384 controls studied, 220 (57.3%) were males and 164 (42.7%) were females, with a ratio of 1.3:1. Mean age was 35 years (range 4 to 62 years).

Of the 298 patients studied, 178 (59.7%) were males and 120 (40.3%) were females with a ratio of 1.4:1. Mean age at diagnosis was 8 years (range 11 months to 18 years). Most of the patients were diagnosed with precursor B cells ALL (pre-B ALL) (95.3%) and the remaining with T cells ALL (T-ALL) (4.7%) (Table 1). *ETV6-RUNX1* was the most frequently observed fusion, whereas the poor prognosis fusions *BCR-ABL1* and those that involved the *KMT2A* gene were detected in few patients. Only 2 patients with T-ALL harbored gene fusions (Table 1). In 14.1% (40/284) of the pre-B ALL patients, it was not possible to obtain the DNA index results; of the pre-B ALL analyzed patients, 16% (39/244) were hyperdiploid (Table 1).

SNPs of *ARID5B* gene associated with ALL susceptibility

The estimated proportion of Native American and European ancestry was not statistically different between controls and patients ($p > 0.05$).

All the SNPs in *ARID5B* were in HWE in both control and patient samples; therefore, the seven SNPs were included in the association studies. The frequency of homozygotes for the risk alleles was significantly increased in patients compared with controls (Table 2). The frequency of the risk alleles was also higher in patients than in controls, but only three SNPs (rs10821936, rs10994982, and rs7089424) showed statistically significant differences (Table 2). The seven SNPs studied were associated with ALL susceptibility (Table 3).

When patients were separated according to B or T lineage, all homozygotes for the risk alleles were significantly more frequent in pre-B ALL patients than in controls. In this group, the frequency of risk alleles of four SNPs (rs10821936, rs10994982, rs7089424, and rs2393732) was significantly higher in patients with pre-B ALL compared with controls. In the pre-B ALL patients, the C allele of the SNP rs4948488 did not reach significance but its p value was at the borderline ($p = 0.0507$) (Table 2). All SNPs were associated with pre-B ALL susceptibility in our population (Table 3). No association between SNPs in *ARID5B* and T-ALL was found, although the frequencies of the risk homozygotes and risk alleles at rs10821936, rs10994982, and rs7089424 were slightly higher in patients with respect to controls (Table 2).

Logistic regression analysis showed that patients who carried two copies of the risk alleles of the SNPs rs10821936,

Table 2 Genotypic and allelic frequencies of the seven SNPs of *ARID5B* gene

SNPs	384 controls *n* (%)	298 patients *n* (%)	p^a	284 pre-B ALL patients *n* (%)	p^a	14 T-ALL patients *n* (%)	p^a
rs10821936							
TT	90 (23.4)	35 (11.7)	< 0.0001*	33 (11.6)	0.0001*	2 (14.3)	0.5366
TC	195 (50.8)	118 (39.6)	0.0041*	110 (38.7)	0.0028*	8 (57.1)	0.7874
CC	99 (25.8)	145 (48.7)	< 0.0001*	141 (49.6)	< 0.0001*	4 (28.6)	0.7628
C risk allele	0.51	0.68	< 0.0001*	0.69	< 0.0001*	0.57	0.7881
rs10994982							
GG	65 (16.9)	31 (10.4)	0.0194*	29 (10.2)	0.0135*	2 (14.3)	1.000
GA	190 (49.5)	109 (36.6)	0.0008*	102 (35.9)	0.0005*	7 (50)	1.000
AA	129 (33.6)	158 (53)	< 0.0001*	153 (53.9)	< 0.0001*	5 (35.7)	1.000
A risk allele	0.58	0.71	0.0004*	0.72	0.0003*	0.61	1.000
rs7089424							
TT	92 (23.9)	36 (12.1)	< 0.0001*	34 (12)	0.0001*	2 (14.3)	0.5343
TG	192 (50)	120 (40.3)	0.0131*	112 (39.4)	0.0075*	8 (57.1)	0.7868
GG	100 (26.1)	142 (47.6)	< 0.0001*	138 (48.6)	0.0001*	4 (28.6)	0.7649
G risk allele	0.51	0.68	< 0.0001*	0.68	0.0001*	0.57	0.7877
rs2393732							
GG	182 (47.4)	123 (41.3)	0.1207	114 (40.1)	0.0700	9 (64.3)	0.2782
GA	173 (45)	128 (42.9)	0.5874	123 (43.3)	0.6938	5 (35.7)	0.5898
AA	29 (7.6)	47 (15.8)	0.0009*	47 (16.5)	0.0005*	0	0.6118
A risk allele	0.30	0.37	0.0596	0.38	0.0313*	0.18	0.5677
rs2393782							
GG	169 (44)	121 (40.6)	0.3909	111 (39.1)	0.2057	10 (71.4)	0.0553
GC	172 (44.8)	125 (41.9)	0.4838	121 (42.6)	0.5819	4 (28.6)	0.2812
CC	43 (11.2)	52 (17.4)	0.0254*	52 (18.3)	0.0101*	0	0.3800
C risk allele	0.34	0.38	0.2268	0.40	0.1040	0.14	0.1576
rs2893881							
AA	170 (44.3)	114 (38.3)	0.1180	107 (37.7)	0.0955	7 (50)	0.7863
AG	169 (44)	128 (42.9)	0.8155	121 (42.6)	0.7523	7 (50)	0.7858
GG	45 (11.7)	56 (18.8)	0.0122*	56 (19.7)	0.0061*	0	0.3836
G risk allele	0.34	0.40	0.0927	0.41	0.0621	0.25	0.7808
rs4948488							
TT	108 (28.1)	68 (22.8)	0.1336	64 (22.5)	0.1078	4 (28.6)	1.000
TC	185 (48.2)	132 (44.3)	0.3158	124 (43.7)	0.2720	8 (57.1)	0.5919
CC	91 (23.7)	98 (32.9)	0.0096*	96 (33.8)	0.0052*	2 (14.3)	0.5353
C risk allele	0.48	0.55	0.0757	0.55	0.0507	0.43	0.7899

[a] *p* values by Fisher's exact test

*Statistically significant *p* value

rs10994982, rs7089424, and rs2393732 had a significantly higher risk for developing childhood ALL and pre-B ALL (OR from 2.33 to 3.85, *p* values from 0.0011 to $2.20e^{-08}$). Particularly, the presence of these SNPs is significantly more associated with susceptibility to childhood ALL and pre-B ALL (Table 3).

Although we found that several risk homozygotes are more frequent in patients than in controls, these SNPs were not associated with the risk of developing ALL in the 40 patients who presented gene fusions (Supplementary data, Table A). Moreover, individual fusions did not present an association with specific alleles. In addition, 22 out of 40 patients (55%) carry at least one of the risk homozygotes of the seven SNPs studied. Fourteen of 22 (63.6%) patients with a gene fusion presented the homozygotes of risk CC (rs10821936), AA (rs10994982), and GG (rs7089424), and 9 of 22 patients (40.9%) with *ETV6-*

Table 3 Association of the SNPs in *ARID5B* with ALL susceptibility in Mexican patients

SNPs	ALL patients			pre-B ALL patients		
	Risk allele frequency	OR (95% CI)	p^a	Risk allele frequency	OR (95% CI)	p^a
rs10821936	0.68			0.69		
T vs C		2.06 (1.64 to 2.57)	$1.72e^{-10}$*		2.11 (1.68 to 2.65)	$7.97e^{-11}$*
TT vs TC		1.55 (0.99 to 2.44)	0.0544		1.53 (0.96 to 2.44)	0.0667
TT vs CC		3.74 (2.34 to 5.96)	$1.33e^{-08}$*		3.85 (2.39 to 6.20)	$1.01e^{-08}$*
rs10994982	0.71			0.72		
G vs A		1.76 (1.40 to 2.22)	$9.15e^{-07}$*		1.81 (1.43 to 2.28)	$4.67e^{-07}$*
GG vs GA		1.20 (0.73 to 1.96)	0.4581		1.20 (0.73 to 1.98)	0.4674
GG vs AA		2.55 (1.56 to 1.53)	0.0001*		2.64 (1.60 to 4.34)	0.0009*
rs7089424	0.68			0.68		
T vs G		2.01 (1.61 to 2.52)	$4.90e^{-10}$*		2.06 (1.64 to 2.59)	$2.44e^{-10}$*
TT vs TG		1.59 (1.02 to 2.49)	0.0395*		1.57 (0.99 to 2.49)	0.0494*
TT vs GG		3.62 (2.28 to 5.76)	$2.20e^{-08}$*		3.73 (2.33 to 5.97)	$1.74e^{-08}$*
rs2393732	0.37			0.38		
G vs A		1.36 (1.08 to 1.71)	0.0075*		1.41 (1.12 to 1.78)	0.0027*
GG vs GA		1.08 (0.78 to 1.50)	0.6063		1.12 (0.81 to 1.56)	0.4704
GG vs AA		2.33 (1.39 to 3.91)	0.0011*		2.51 (1.49 to 4.23)	0.0004*
rs2393782	0.38			0.40		
G vs C		1.23 (0.98 to 1.54)	0.0648		1.29 (1.03 to 1.62)	0.0235*
GG vs GC		1.01 (0.73 to 1.40)	0.9289		1.07 (0.76 to 1.49)	0.6868
GG vs CC		1.68 (1.05 to 2.69)	0.0269*		1.84 (1.15 to 2.94)	0.0103*
rs2893881	0.40			0.41		
A vs G		1.32 (1.06 to 1.65)	0.0128*		1.36 (1.09 to 1.71)	0.0062*
AA vs AG		1.12 (0.81 to 1.57)	0.4698		1.13 (0.81 to 1.59)	0.4523
AA vs GG		1.85 (1.17 to 2.93)	0.0078*		1.97 (1.24 to 3.13)	0.0034*
rs4948488	0.55			0.55		
T vs C		1.32 (0.71 to 1.64)	0.0093*		1.36 (1.09 to 1.69)	0.0054*
TT vs TC		1.13 (0.77 to 1.65)	0.5151		1.13 (0.77 to 1.69)	0.5292
TT vs CC		1.69 (1.11 to 2.56)	0.0130*		1.76 (1.15 to 2.68)	0.0083*

[a] *p* values by Fisher's exact test

*Statistically significant *p* value

RUNX1, *E2A-PBX1*, *KMT2A-MLLT3*, *KMT2A-EPS15*, and *KMT2A-AFF1* were positive to the homozygotes of risk CC (rs10821936), AA (rs10994982), GG (rs7089424), and CC (rs4948488). Interestingly, one patient with *ETV6-RUNX1* was positive for all the homozygotes of risk analyzed.

In contrast, all SNPs conferred risk for developing pre-B ALL in the 39 patients with hyperdiploidy (Supplementary data, Tables A and B). In this study, we found a positive association between the SNPs rs2393732, rs2393782, and rs4948488 with childhood pre-B hyperdiploid ALL.

We also analyzed our population based on the distribution of patients according to the age at diagnosis, and 3 age peaks of incidence were found (Supplementary data, Fig. B,). The first and main peak occurs between 3 and 7 years of age and includes 127 (44.7%) patients. The second peak occurs between 8 and 10 years of age and comprises 45 (15.8%) patients. A third peak occurs between 11 and 14 years of age and includes 63 (22.2%) patients (Supplementary data, Table A). All SNPs of *ARID5B* were associated with the risk of developing pre-B ALL in the first group, 4 out of seven SNPs (rs10821936, rs10994982, rs7089424, and rs4948488) were associated with pre-B ALL susceptibility in the second group, and the same SNPs except rs4948488 were associated with the susceptibility to develop the disease in the third group (Supplementary data, Table B).

All SNPs of *ARID5B* gene were subjected to haplotype analysis to determine their association with childhood ALL. The analysis revealed a strong linkage disequilibrium (LD) ($D' = 0.96$) between rs4948488 and rs2893881 (block 1), this block covered 3 kb and carried the risk alleles C and G.

SNPs rs10821936, rs10994982, and rs7089424 also were in LD (D′ = 0.96) and formed block 2, which covered 42 kb and carried the risk alleles C, A, and G. Only the CAG haplotype showed a statistically significant association with ALL risk ($p < 0.00001$) after a 100,000 permutation test; this haplotype was present in 57.3% of our patients.

Discussion

The ancestry analysis in our patients and controls did not reveal statistical differences; therefore, the possibility of spurious associations by population stratification was diminished.

In this study, we observed that frequencies of the risk alleles of the seven SNPs of *ARID5B* gene are higher in Mexican controls and patients than those documented in other populations (Supplementary data, Figs. C and D respectively). For example, the frequency of the C risk allele at rs10821936 in ALL children is lower in Caucasian (0.44 to 0.48) [11, 39], African-American (0.32 to 0.33) [15, 40, 41], Chinese (0.45) [38], Indian (0.61) [5], and Hispanic-American (0.59 to 0.63) [39, 40], than in Mexicans (0.68); and it is slightly lower than that reported in Guatemalan (0.73) patients [2]. Interestingly, the frequency of the C allele in healthy Zapotec population from Mexico is very close to that found in our controls (0.52 vs 0.51 respectively) [39]. These differences are fundamental in the genetic predisposition, since the Amerindian component could influence the high frequency of the risk alleles in our Mexican mestizo population.

Recently, it has been described that the CGAACACAA haplotype formed by the SNPs rs6479778, rs6479779, rs7073837, rs10994982, rs10740055, rs7923074, rs10821936, rs7896246, and rs10821938 of *ARID5B* can contribute to the increased risk of ALL in Yemeni population [1]. It is possible that we have a similar behavior, because each SNP of *ARID5B* confers an individual effect on the risk for developing the disease, and the presence of the CAG (rs10821936, rs10994982, and rs7089424) haplotype is strongly associated with ALL susceptibility. Therefore, this could play an important role in increasing the risk of childhood ALL and also could contribute to the higher incidence of ALL in our population.

There is no direct evidence of the biological effect of the SNPs rs10821936, rs10994982, rs7089424, rs2393732, rs2393782, rs2893881, and rs4948488 on the function of *ARID5B*, which limits establishing the real impact of the presence of these SNPs with the development of pre-B ALL. Recent studies indicate that risk alleles at rs7090445 cause low expression of *ARID5B* inducing abnormal growth and development of lymphoid cells by histone demethylation, which produces inactivation of tumor suppressor genes as *p53* [12, 19, 33]. Similar to rs7090445, the SNPs rs10821936 and rs7089424 are located in intron 3 (GRCh38.p7); therefore, these SNPs could be reducing the expression of *ARID5B*. In order to know if the risk alleles of the SNPs evaluated in this study produce a similar effect as rs7090445 in *ARID5B*, an expression analysis must be performed in our patients.

The risk allele of the SNP rs2393782 was not significantly higher in our ALL and pre-B ALL studied population (p values of 0.2268 and 0.1040 respectively), although it was significantly associated in Hispanics from USA ($p = 0.0290$) [39]; we must emphasize that the sample size analyzed in both studies was not very different (330 Hispanic children and 298 Mexican children). This suggests that each population has a genetic background with specific genetic variations such as SNPs, which could be influencing the susceptibility or protection for the development of diverse diseases, including ALL [4].

For us, it is essential to understand the relationship between the SNPs of *ARID5B* and the susceptibility to develop childhood ALL, since there is limited information regarding Mexican children. Although Mexicans are Hispanic, the proportion of European, Native American, and African ancestry depends on the geographical region and is very different between them [9]. This ancestry differences could thus contribute to the increased risk of susceptibility and could also explain the high incidence of ALL in Mexicans. It is important to mention that all controls and patients included were Mexican mestizos with a high proportion of Amerindian ancestry (51 to 54%) in contrast to Colombians (19%), and Puerto Ricans (15%), and a low proportion of African ancestry (5%), compared with Colombians (10%) and Puerto Ricans (18%) [31, 34]. These differences are fundamental in the genetic predisposition, since some authors have proposed that Amerindian ancestry may represent a potential risk factor for developing ALL and could be in part the basis of the higher risk of relapse in Native Americans and Hispanics [2, 8, 37]. On the other hand, the C risk allele at rs10821936 presented a high prevalence in healthy indigenous population as Zapotecs (0.52), Mixa (0.64), Mixe (0.83), Kaqchikel (0.72), and Kiche (0.76), and its frequency has a positive association with Native American ancestry [39].

Similar to other authors, we did not find any association between SNPs of *ARID5B* and T-ALL susceptibility [21, 26, 35, 38]; however, the number of T-ALL patients studied was small (14 of 298 patients). We did not find an association between the SNPs and gene fusions including *ETV6-RUNX1* and *TCF3-PBX1* neither [39].

To our knowledge, this is the first report showing the positive association between SNPs rs2393732, rs2393782, and rs4948488 with the risk to childhood B-hyperdiploid ALL. Based on previous reports, we showed that the frequency of all risk alleles was higher in Mexican pre-B ALL patients with hyperdiploid with respect to other populations [21, 26, 28, 39, 41] (Supplementary data, Fig. A). In our population, the presence of G allele at rs2893881 conferred a greater risk of

developing pre-B-hyperdiploid ALL (OR, 2.29) than in Hispanic children (OR, 1.71) [16]. The SNPs rs10821936, rs10994982, and rs7089424 have been previously associated with high susceptibility to pre-B-hyperdiploid ALL subtype in pediatric White, Black, Hispanic, French-Canadian [8, 15, 21, 26, 28, 39, 41], and now in Mexican patients.

Considering the age peak of incidence, our first peak is similar to the 4–7-year peak documented by Rivera-Luna (personal communication) in 5861 patients with ALL from the Popular Medical Insurance; this is a program which provides oncological treatment to children with cancer in Mexico. The age range of the second peak is the same reported by Pérez-Saldivar [29]. Interestingly, this study suggests that rs2393732, rs2393782, rs2893881, and rs4948488 do not confer pre-B ALL susceptibility in older children and adolescent patients.

Our findings suggest that the genetic background of the Mexican population could be positively influencing the susceptibility to ALL development, specifically pre-B ALL, and inherited genetic variants such as SNPs of *ARID5B* could increase the risk for the disease. In this case, the Native American ancestry could be related to high frequencies of the risk alleles.

In Mexican children, the SNPs rs10821936, rs10994982, rs7089424, and rs2393732 of *ARID5B* gene are significantly more associated with an increased risk for developing childhood ALL, specifically pre-B ALL. The frequency of all risk alleles was higher than those reported in the Hispanic population from the USA. This could partially explain the high incidence of childhood ALL in Mexico.

Acknowledgments We would like to thank the participating patients, without whom this study would not have been possible, and the nurses of the Ambulatory Care Unit (AQUA) at the Instituto Nacional de Pediatria, for their assistance.

Funding This work was supported by the grants from Fondos del Presupuesto Federal para la Investigación (project 085/2012), Consejo Nacional de Ciencia y Tecnologia (CONACyT) Desarrollo Cientifico para Atender Problemas Nacionales (project 216163), and in part by the Intramural Program of the National Cancer Institute.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This study was reviewed and approved by the Institutional Research and Ethics Committees from both participant Institutions in accordance with the ethical principles of the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent All volunteers, parents, or legal tutors were previously informed about the study, and before biological samples were collected, they provided a signed, written informed consent letter to participate.

References

1. Al-absi B, Razif MFM, Noor SM et al (2017) Contributions of *IKZF1* , *DDC* , *CDKN2A* , *CEBPE* , and *LMO1* gene polymorphisms to acute lymphoblastic leukemia in a Yemeni population. Genet Test Mol Biomarkers 21:592–599. https://doi.org/10.1089/gtmb.2017.0084
2. Archer NP, Perez-Andreu V, Stoltze U, Scheurer ME, Wilkinson AV, Lin TN, Qian M, Goodings C, Swartz MD, Ranjit N, Rabin KR, Peckham-Gregory EC, Plon SE, de Alarcon PA, Zabriskie RC, Antillon-Klussmann F, Najera CR, Yang JJ, Lupo PJ (2017) Family-based exome-wide association study of childhood acute lymphoblastic leukemia among Hispanics confirms role of ARID5B in susceptibility. PLoS One 12:1–12. https://doi.org/10.1371/journal.pone.0180488
3. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265. https://doi.org/10.1093/bioinformatics/bth457
4. Bell JI (2002) Single nucleotide polymorphisms and disease gene mapping. Arthritis Res 4:273–278. https://doi.org/10.1186/ar555
5. Bhandari P, Ahmad F, Mandava S, Das BR (2016) Association of genetic variants in ARID5B, IKZF1 and CEBPE with risk of childhood de novo B-lineage acute lymphoblastic leukemia in India. Asian Pac J Cancer Prev 17:3987–3993
6. Bhatia S, Sather HN, Heerema NA et al (2002) Racial and ethnic differences in survival of children with acute lymphoblastic leukemia. Blood 100:1957–1964. https://doi.org/10.1182/blood-2002-02-0395
7. Buffler PA, Kwan ML, Reynolds P, Urayama KY (2005) Environmental and genetic risk factors for childhood leukemia: appraising the evidence. Cancer Investig 23:60–75. https://doi.org/10.1081/CNV-46402
8. Chokkalingam AP, Hsu LI, Metayer C, Hansen HM, Month SR, Barcellos LF, Wiemels JL, Buffler PA (2013) Genetic variants in ARID5B and CEBPE are childhood ALL susceptibility loci in Hispanics. Cancer Causes Control 24:1789–1795. https://doi.org/10.1007/s10552-013-0256-3
9. Conomos MP, Laurie CA, Stilp AM, Gogarten SM, McHugh CP, Nelson SC, Sofer T, Fernández-Rhodes L, Justice AE, Graff M, Young KL, Seyerle AA, Avery CL, Taylor KD, Rotter JI, Talavera GA, Daviglus ML, Wassertheil-Smoller S, Schneiderman N, Heiss G, Kaplan RC, Franceschini N, Reiner AP, Shaffer JR, Barr RG, Kerr KF, Browning SR, Browning BL, Weir BS, Avilés-Santa ML, Papanicolaou GJ, Lumley T, Szpiro AA, North KE, Rice K, Thornton TA, Laurie CC (2016) Genetic diversity and association studies in US Hispanic/Latino populations: applications in the Hispanic community health study/study of Latinos. Am J Hum Genet 98:165–184. https://doi.org/10.1016/j.ajhg.2015.12.001
10. Dickinson HO, Hodgson JT, Parker L (2003) Comparison of Health and Safety Executive and Cumbrian birth cohort studies of risk of leukaemia/non-Hodgkin's lymphoma in relation to paternal preconceptional irradiation. J Radiol Prot 23:385–403. https://doi.org/10.1088/0952-4746/23/4/003
11. Evans TJ, Milne E, Anderson D, de Klerk NH, Jamieson SE, Talseth-Palmer BA, Bowden NA, Holliday EG, Rudant J, Orsi L, Richardson E, Lavis L, Catchpoole D, Attia JR, Armstrong BK, Clavel J, Scott RJ (2014) Confirmation of childhood acute lymphoblastic leukemia variants, ARID5B and IKZF1, and interaction with parental environmental exposures. PLoS One 9:e110255. https://doi.org/10.1371/journal.pone.0110255
12. Ge Z, Han Q, Gu Y, Ge Q, Ma J, Sloane J, Gao G, Payne KJ, Szekely L, Song C, Dovat S (2018) Aberrant ARID5B expression and its association with Ikaros dysfunction in acute lymphoblastic leukemia. Oncogenesis 7:84. https://doi.org/10.1038/s41389-018-0095-x
13. Gharbi H, Ben Hassine I, Soltani I, Safra I, Ouerhani S, Bel Haj Othmen H, Teber M, Farah A, Amouri H, Toumi NH, Abdennebi S, Abbes S, Menif S (2016) Association of genetic variation in IKZF1,

ARID5B, CDKN2A, and CEBPE with the risk of acute lymphoblastic leukemia in Tunisian children and their contribution to racial differences in leukemia incidence. Pediatr Hematol Oncol 33:157–167. https://doi.org/10.3109/08880018.2016.1161685

14. Gutiérrez-Camino A, López-López E, Martín-Guerrero I, Sánchez-Toledo J, García de Andoin N, Carboné Bañeres A, García-Miguel P, Navajas A, García-Orad Á (2013) Intron 3 of the ARID5B gene: a hot spot for acute lymphoblastic leukemia susceptibility. J Cancer Res Clin Oncol 139:1879–1886. https://doi.org/10.1007/s00432-013-1512-3
15. Healy J, Richer C, Bourgey M, Kritikou EA, Sinnett D (2010) Replication analysis confirms the association of ARID5B with childhood B-cell acute lymphoblastic leukemia. Haematologica 95:1608–1611. https://doi.org/10.3324/haematol.2010.022459
16. Hsu LI, Chokkalingam AP, Briggs FBS, Walsh K, Crouse V, Fu C, Metayer C, Wiemels JL, Barcellos LF, Buffler PA (2015) Association of genetic variation in IKZF1, ARID5B, and CEBPE and surrogates for early-life infections with the risk of acute lymphoblastic leukemia in Hispanic children. Cancer Causes Control 26:609–619. https://doi.org/10.1007/s10552-015-0550-3
17. Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x
18. Kadan-Lottick NS, Ness KK, Bhatia S, Gurney JG (2003) Survival variability by race and ethnicity in childhood acute lymphoblastic leukemia. J Am Med Assoc 290:2008–2014. https://doi.org/10.1001/jama.290.15.2008
19. Lee KH, Park JW, Sung HS, Choi YJ, Kim WH, Lee HS, Chung HJ, Shin HW, Cho CH, Kim TY, Li SH, Youn HD, Kim SJ, Chun YS (2015) PHF2 histone demethylase acts as a tumor suppressor in association with p53 in cancer. Oncogene 34:2897–2909. https://doi.org/10.1038/onc.2014.219
20. Lin CY, Li MJ, Chang JG, Liu SC, Weng T, Wu KH, Yang SF, Huang FK, Lo WY, Peng CT (2014) High-resolution melting analyses for genetic variants in ARID5B and IKZF1 with childhood acute lymphoblastic leukemia susceptibility loci in Taiwan. Blood Cells Mol Dis 52:140–145. https://doi.org/10.1016/j.bcmd.2013.10.003
21. Linabery AM, Blommer CN, Spector LG, Davies SM, Robison LL, Ross JA (2013) ARID5B and IKZF1 variants, selected demographic factors, and childhood acute lymphoblastic leukemia: a report from the Children's oncology group. Leuk Res 37:936–942. https://doi.org/10.1016/j.leukres.2013.04.022
22. Martinez-Marignac VL, Valladares A, Cameron E, Chan A, Perera A, Globus-Goldberg R, Wacher N, Kumate J, McKeigue P, O'Donnell D, Shriver MD, Cruz M, Parra EJ (2007) Admixture in Mexico City: implications for admixture mapping of type 2 diabetes genetic risk factors. Hum Genet 120:807–819. https://doi.org/10.1007/s00439-006-0273-3
23. Matasar MJ, Ritchie EK, Consedine N, Magai C, Neugut AI (2006) Incidence rates of the major leukemia subtypes among U.S. Hispanics, Blacks, and non-Hispanic Whites. Leuk Lymphoma 47:2365–2370. https://doi.org/10.1080/10428190600799888
24. McNeil DE, Coté TR, Clegg L, Mauer A (2002) SEER update of incidence and trends in pediatric malignancies: acute lymphoblastic leukemia. Med Pediatr Oncol 39:554–557. https://doi.org/10.1002/mpo.10161
25. Migliorini G, Fiege B, Hosking FJ, Ma Y, Kumar R, Sherborne AL, da Silva Filho MI, Vijayakrishnan J, Koehler R, Thomsen H, Irving JA, Allan JM, Lightfoot T, Roman E, Kinsey SE, Sheridan E, Thompson P, Hoffmann P, Nothen MM, Muhleisen TW, Eisele L, Zimmermann M, Bartram CR, Schrappe M, Greaves M, Stanulla M, Hemminki K, Houlston RS (2013) Variation at 10p12.2 and 10p14 influences risk of childhood B-cell acute lymphoblastic leukemia and phenotype. Blood 122:3298–3307. https://doi.org/10.1182/blood-2013-03-491316
26. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, Kinsey SE, Lightfoot T, Roman E, Irving JAE, Allan JM, Tomlinson IP, Taylor M, Greaves M, Houlston RS (2009) Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet 41:1006–1010. https://doi.org/10.1038/ng.430
27. Pastorczak A, Górniak P, Sherborne A, Hosking F, Trelińska J, Lejman M, Szczepański T, Borowiec M, Fendler W, Kowalczyk J, Houlston RS, Młynarski W (2011) Role of 657del5 NBN mutation and 7p12.2 (IKZF1), 9p21 (CDKN2A), 10q21.2 (ARID5B) and 14q11.2 (CEBPE) variation and risk of childhood ALL in the Polish population. Leuk Res 35:1534–1536. https://doi.org/10.1016/j.leukres.2011.07.034
28. Paulsson K, Forestier E, Lilljebjorn H, Heldrup J, Behrendtz M, Young BD, Johansson B (2010) Genetic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. Proc Natl Acad Sci 107:21719–21724. https://doi.org/10.1073/pnas.1006981107
29. Pérez-Saldivar ML, Fajardo-Gutiérrez A, Bernáldez-Ríos R et al (2011) Childhood acute leukemias are frequent in Mexico City: descriptive epidemiology. BMC Cancer 11:355. https://doi.org/10.1186/1471-2407-11-335
30. Prasad RB, Hosking FJ, Vijayakrishnan J, Papaemmanuil E, Koehler R, Greaves M, Sheridan E, Gast A, Kinsey SE, Lightfoot T, Roman E, Taylor M, Pritchard-Jones K, Stanulla M, Schrappe M, Bartram CR, Houlston RS, Kumar R, Hemminki K (2010) Verification of the susceptibility loci on 7p12.2, 10q21.2, and 14q11.2 in precursor B-cell acute lymphoblastic leukemia of childhood. Blood 115:1765–1767. https://doi.org/10.1182/blood-2009-09-241513
31. Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, McDonald GJ, Tandon A, Schirmer C, Neubauer J, Bedoya G, Duque C, Villegas A, Bortolini MC, Salzano FM, Gallo C, Mazzotti G, Tello-Ruiz M, Riba L, Aguilar-Salinas CA, Canizales-Quinteros S, Menjivar M, Klitz W, Henderson B, Haiman CA, Winkler C, Tusie-Luna T, Ruiz-Linares A, Reich D (2007) A Genomewide admixture map for Latino populations. Am J Hum Genet 80:1024–1036. https://doi.org/10.1086/518313
32. Rivera-Luna R, Velasco-Hidalgo L, Zapata-Tarrés M, Cárdenas-Cardos R, Aguilar-Ortiz MR (2017) Current outlook of childhood cancer epidemiology in a middle-income country under a public health insurance program. Pediatr Hematol Oncol 34:43–50. https://doi.org/10.1080/08880018.2016.1276236
33. Studd JB, Vijayakrishnan J, Yang M, Migliorini G, Paulsson K, Houlston RS (2017) Genetic and regulatory mechanism of susceptibility to high-hyperdiploid acute lymphoblastic leukaemia at 10p21.2. Nat Commun 8:1–9. https://doi.org/10.1038/ncomms14616
34. Tang H, Choudhry S, Mei R, Morgan M, Rodriguez-Cintron W, Burchard EG, Risch NJ (2007) Recent genetic selection in the ancestral admixture of Puerto Ricans. Am J Hum Genet 81:626–633. https://doi.org/10.1086/520769
35. Treviño LR, Yang W, French D et al (2009) Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet 41:1001–1005. https://doi.org/10.1038/ng.432
36. Vijayakrishnan J, Sherborne AL, Sawangpanich R, Hongeng S, Houlston RS, Pakakasama S (2010) Variation at 7p12.2 and 10q21.2 influences childhood acute lymphoblastic leukemia risk in the Thai population and may contribute to racial differences in leukemia incidence. Leuk Lymphoma 51:1870–1874. https://doi.org/10.3109/10428194.2010.511356
37. Walsh KM, Chokkalingam AP, Hsu LI, Metayer C, de Smith AJ, Jacobs DI, Dahl GV, Loh ML, Smirnov IV, Bartley K, Ma X, Wiencke JK, Barcellos LF, Wiemels JL, Buffler PA (2013) Associations between genome-wide Native American ancestry, known risk alleles and B-cell ALL risk in Hispanic children. Leukemia 27:2416–2419. https://doi.org/10.1038/leu.2013.130
38. Wang Y, Chen J, Li J, Deng J, Rui Y, Lu Q, Wang M, Tong N, Zhang Z, Fang Y (2013) Association of three polymorphisms in

ARID5B, IKZF1and CEBPE with the risk of childhood acute lymphoblastic leukemia in a Chinese population. Gene 524:203–207. https://doi.org/10.1016/j.gene.2013.04.028

39. Xu H, Cheng C, Devidas M, Pei D, Fan Y, Yang W, Neale G, Scheet P, Burchard EG, Torgerson DG, Eng C, Dean M, Antillon F, Winick NJ, Martin PL, Willman CL, Camitta BM, Reaman GH, Carroll WL, Loh M, Evans WE, Pui CH, Hunger SP, Relling MV, Yang JJ (2012) ARID5B genetic polymorphisms contribute to racial disparities in the incidence and treatment outcome of childhood acute lymphoblastic leukemia. J Clin Oncol 30:751–757. https://doi.org/10.1200/JCO.2011.38.0345
40. Xu H, Yang W, Perez-Andreu V, Devidas M, Fan Y, Cheng C, Pei D, Scheet P, Burchard EG, Eng C, Huntsman S, Torgerson DG, Dean M, Winick NJ, Martin PL, Camitta BM, Bowman WP, Willman CL, Carroll WL, Mullighan CG, Bhojwani D, Hunger SP, Pui CH, Evans WE, Relling MV, Loh ML, Yang JJ (2013) Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. J Natl Cancer Inst 105:733–742. https://doi.org/10.1093/jnci/djt042
41. Yang W, Trevĩo LR, Yang JJ et al (2010) ARID5B SNP rs10821936 is associated with risk of childhood acute lymphoblastic leukemia in blacks and contributes to racial differences in leukemia incidence. Leukemia 24:894–896. https://doi.org/10.1038/leu.2009.277