Published in final edited form as:

Ann Hematol. 2019 October ; 98(10): 2379–2388. doi:10.1007/s00277-019-03730-x.

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 Lopez-Ruiz**[5], **Monica Anabell Malavar-Guadarrama**[6], **Rocío Cárdenas-Cardós**[3], **Rogelio Paredes-Aguilera**[4], **Roberto Rivera-Luna**[3], **Michael Dean**[7], **Patricia Pérez-Vera**[1]

[1]Cancer and Genetics Laboratory, National Institute of Pediatrics, Mexico City, Mexico.

[2]Molecular Biology Laboratory, National Institute of Pediatrics, Mexico City, Mexico.

[3]Department of Oncology, National Institute of Pediatrics, Mexico City, Mexico.

[4]Department of Hematology, National Institute of Pediatrics, Mexico City, Mexico.

[5]Department of Oncology, Hospital of Pediatric Specialities from Tuxtla Gutierrez, Chiapas, Mexico,

[6]Department of Hematology, Hospital of Pediatric Specialities from Tuxtla Gutierrez, Chiapas, Mexico,

[7]Laboratory of Translational Genomics, Division of Cancer Epidemiolgy & Genetics, National Cancer Institute, Bethesda, Maryland, USA.

SUMMARY

Background: In Mexico acute lymphoblastic leukemia (ALL) is the most common cancer in children. Recently, several SNPs of *ARID5B* have been associated with ALL susceptibility and are more strongly related with ALL risk in Hispanics. However, these observations are not representative of the situation in Mexico since the proportion of Mexican patients included is unknown. The purpose of this study was to determine the association between the SNPs rs10821936, rs10994982, rs7089424, rs2393732, rs2393782, rs2893881, rs4948488 of *ARID5B* with ALL susceptibility in Mexican children.

Methods: The study included 384 controls and 298 children with ALL recruited at National Institute of Pediatrics and Hospital of Pediatric Specialities from Tuxtla Gutierrez, Chiapas. Genotyping analysis was done with genomic DNA and pre-designed TaqMan probes. Genotypic and allelic frequencies were calculated to compare the differences between controls and patients (Fisher's exact test). The odds ratio (OR) was calculated to determine the association between SNPs and ALL susceptibility. Haplotype and ancestry analysis was done (Haploview and STRUCTURE programs).

Findings: The SNPs rs10821936, rs10994982 and rs7089424 were associated with ALL and B-ALL susceptibility, and rs2393732 was also associated with B-ALL. No association with T-ALL was found. The CAG haplotype (rs10821936, rs10994982 and rs7089424) was stronger associated

with ALL risk in our population. The frequency of the risk alleles of the 7 SNPs of *ARID5B* gene was higher than in other populations.

Interpretation: The SNPs rs10821936, rs10994982, rs7089424 and rs2393732 were significantly associated with an increased risk to develop childhood ALL, specifically B-ALL. The genetic background of our population could be influencing the susceptibility to ALL and could also explain in part its high incidence in Mexico.

Funding: This work was supported by Fondos del Presupuesto Federal para la Investigación and Consejo Nacional de Ciencia y Tecnologia (CONACyT).

INTRODUCTION

Acute lymphoblastic leukemia (ALL) is characterized by the uncontrolled proliferation and malignant transformation of lymphoid progenitor cells in the bone marrow, blood and extramedullary sites. It is the most common subtype of leukemia diagnosed in children under 18 years.[1]

The incidence of childhood ALL in the United States of America (USA) varies according to race or ethnic group, for example in Hispanic children it 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.[2,3] 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.[4]

ALL is a multifactorial disease, in which environmental (prenatal exposure to X-rays, pesticide 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 develop of the disease;[5,6] however these factors are not enough to explain the high incidence of childhood ALL. 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.[7–17] Apparently the treatment outcome, prognosis and overall survival of childhood ALL depends on race (Asian, Black, White, Native American) or ethnicity (Hispanic and Non-Hispanic).[18,19]

Interestingly, specific SNPs of *ARID5B* (rs7089424, rs10994982, rs10740055, rs10821936, rs2393782, rs17215180, rs2393782), *CDKN2A* (rs3218018, rs3731217), *IKZF1* (rs4132601, rs7780012), *CEBPE* (rs2239633) and *PIP4K2A* (rs7088318), are more associated with ALL risk in Hispanic than in White children.[20–22] Furthermore, some of them (rs7089424, rs7088318, rs2239633, rs3731217) are highly correlated with the increase of Native American ancestry.[23] 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 between populations.[21]

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* and the susceptibility to develop ALL in Mexican children.

METHODS

Study population:

A total of 298 patients younger than 18 years with 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 Hematology and Oncology Services of the National Institute of Pediatrics in Mexico City, and the Hospital of Pediatric Specialities from Tuxtla Gutierrez, Chiapas, Mexico. The diagnosis was established by cytomorphology, immunophenotyping and molecular biology for the most common translocations. The blast lineage was determined by the expression of different CD cell surface markers; expression of CD10, CD19, CD20 and CD22 defines 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 volunters 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 volunters, or with parents or legal tutors for individuals younger than 18 years. None elegibility criteria was used to select the participants.

Patients and controls were Mexican mestizos 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 to the ethical principles of the Declaration of Helsinki. Volunters, 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.

Biological samples and genomic DNA extraction

Saliva samples were collected from all patients according to a commercial kit intructions (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, Vermont, USA).

ARID5B genotyping

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 seconds and 95°C for 10 minutes; followed by amplification step for 40 cycles of 15 seconds at 95°C and 60 seconds at 60°C; final step consisted for 30 seconds at 60°C. Each experiment included one negative control and three positive controls (allele 1 homozygous, heterozygous and allele 2 homozygous). Randomly repeated about 30% samples for replication and concordance was absolute.

Ancestry analysis

Ten ancestry informative markers (AIMs) (rs2695, rs2862, rs3340, rs17203, rs203096, rs223830, rs722098, rs1800498, rs2065160, rs2814778) was 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 (δ>0.44) in Mexican population.[24] For the control group, the ancestry analysis was carried out with the genotypes obtained of 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.[25] The parameters of the analysis were as follow: 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 t-student test, and p values less than 0.05 was considered statistically significant.

Statistical analysis

Seven 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 was determined by *p* value, and the odds ratios (ORs) with 95% confidence intervals (95% IC) were also obtained (DeFinetti software; https://ihg.gsf.de/cgi-bin/hw/hwa2.pl).

All SNPs were submitted to haplotype analysis using the Haploview program[26] to known if they were in linkage disequilibrium (LD) (D'>0.80); the haplotypes blocks formed were submited to association analysis with susceptibility of ALL, the *p* values were obtained by permutation analysis (100,000 repetitions). A positive association to ALL susceptibility was considered with a *p* value <0.05.

RESULTS

Characteristics of controls and patients

Of the controls studied, 220 (57·3%) were male and 164 (42·7%) female. Mean age was 35 years (range 4 to 62 years).

Of the patients studied, 178 (59·7%) were male and 120 (40·3%) female. Mean age at diagnosis was 8 years (range 11 months to 18 years). Most of the patients were diagnosed with B-ALL (95·3%) and the remaining had 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 *KMT2A*, were detected in a few patients (Table 1).

Association between the SNPs of ARID5B and ALL susceptibility

The estimated proportion of Native American and European ancestry was not statistically different between controls and patients.

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 allele was significantly increased in patients compared to controls. The frequency of the risk alleles was also higher in patients than in controls, but only three SNPs showed statistically significant diferrences (Table 2).

When patients were separated according to B or T lineage, all homozygotes for the risk allele were significantly more frequent in B-ALL than in ALL cases and controls. For four SNPs, the frequency of risk alleles was significantly higher in patients with B-ALL than in controls (Table 2). 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 respect to controls (Table 2).

Logistic regression analysis showed that patients who carried two copies of the risk alleles of the seven SNPs had a significantly higher risk for developing childhood ALL and B-ALL. Particularly, the presence of the risk allele of rs10821936, rs10994982, rs7089424, and

rs2393732 is signifcantly associated with susceptibility to childhood ALL and B-ALL (Table 3).

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; risk alleles C and G), it covers 3 kb. SNPs rs10821936, rs10994982 and rs7089424 also were in LD (D'=0.96) (block 2; risk alleles C, A and G), it covers 42 kb. Only the CAG haplotype showed a statistically significant association with ALL risk ($p<0.00001$) after a 100,000 permutation test, and this haplotype was present in 57.3% of our patients.

DISCUSSION

The ancestry analysis in our patients and controls did not revealed statistical differences, therefore the posibility of spuriuos associations by population stratification was diminished.

In this study we observed that the frequencies of the risk alleles of the seven SNPs of *ARID5B* gene are higher than those documented in other populations, for example: the frequency of the C risk allele at rs10821936 is lower in Caucasians (0.44 to 0.48),[7,21,27] African-Americans (0.32 to 0.33),[12,28] Chinese (0.45),[13] Indians (0.61),[17] and Hispanic-Americans (0.59 to 0.63),[12,21] than in Mexicans (0.68) ALL children; and it is slightly lower than that reported in Guatemalan (0.73) patients.[29] Interestingly, the frequency of the C allele in healthy Zapotec population from Mexico is very closer to that found in our controls (0.52 *vs* 0.51 respectively).[21]

Recently, have been describe 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;[30] is posible that we have a similar behavior, because each SNP of *ARID5B* confers a individual effect on the risk for developing the disease, and the presence of the CAG haplotype is stronger associated with ALL susceptibility; so this could plays an important role to increase the risk of childhood ALL, and also contribute to the higher incidence of LLA in our population.

The risk allele of the SNP rs2393782 did not confer susceptibility to ALL in Mexican patients, although in Hispanics it was significantly associated;[21] we must emphasize that the sample size analyzed in both studies was not very different (330 Hispanics and 298 Mexicans). This suggest that each population has a genetic background with specific genetic variations such as SNPs, which could be influencing the susceptiblity or protection for the development of diverse diseases, including ALL.[31] For us, it is essential to understand the relationship between the SNPs of *ARID5B* and the susceptibility to develop childhood ALL, since there is little information for Mexicans. Although Mexicans are Hispanic, the proportion of European, Amerindian and African ancestry is very different,[32] and it could contribute to increase the risk of susceptibility and could explain the high incidence; 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 (20%)

and Puerto Ricans (16 to 20%), and low proportion of African ancestry (5%) respect Colombians (10%) and Puerto Ricans (20%).[33] These differences are fundamental in the genetic predisposition, since some authors have proposed that Amerindian ancestry may represent a potential risk factor to develop ALL and could be in part the basis of the higher risk of relapse in Native Americans and Hispanics.[20,23,29] On the other hand, the C risk allele at rs10821936 has 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 had a positive association with Amerindian ancestry.[21]

Similar to other authors, we do not find any association between SNPs of *ARID5B* and T-ALL susceptibility,[7,8,13,34] however, the number of patients studied was small (14 of 298 patients).

Our findings suggest that the genetic background of the Mexican population could be influencing the susceptibility to ALL development specifically B-ALL, and inherited genetic variants such as SNPs of *ARID5B* could increase the risk of disease. In conclusion, the SNPs rs10821936, rs10994982, rs7089424 and rs2393732 of *ARID5B* gene are significantly associated with an increased risk to develop childhood ALL, specifically B-ALL. The frequency of all risk alleles was higher than those reported in other populations including Hispanics. The genetic background of our population could be positively influencing the susceptibility to ALL and could also explain in part its high incidence in Mexico.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

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

This work was supported by the grants from Fondos del Presupuesto Federal para la Investigación (project 085/2012) and 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.

REFERENCES

1. McNeil DE, Cote TR, Clegg L, Mauer A. SEER update of incidence and trends in pediatric malignancies: acute lymphoblastic leukemia. Med Pediatr Oncol. 2002; 39(6):554–557. [PubMed: 12376977]
2. Matasar MJ, Ritchie EK, Consedine N, Magai C, Neugut AI. Incidence rates of the major leukemia subtypes among US Hispanics, Blacks, and non-Hispanic Whites. Leuk Lymphoma. 2006; 47(11):2365–70. [PubMed: 17107911]
3. Yang W, Treviño LR, Yang JJ, et al. *ARID5B* SNP rs10821936 is associated with risk of childhood acute lymphoblastic leukemia in blacks and contributes to racial differences in leukemia incidence. Leukemia. 2010; 24(4):894–6. [PubMed: 20054350]
4. Rivera-Luna R, Velasco-Hidalgo L, Zapata-Tarrés M, Cárdenas-Cardós R, Aguilar-Ortiz MR. Current outlook of childhood cancer epidemiology in a middle-income country under a public health insurance program. Pediatr Hematol Oncol. 2017; 34(1):43–50. [PubMed: 28287336]

5. Dickinson HO, Hodgson JT, Parker L. Comparison of health and safety executive and cumbrian birth cohort studies of risk of leukaemia/non-Hodkingś lymphoma in relation to paternal preconceptional irradiation. J Radiol Prot. 2003; 23(4):385–403. [PubMed: 14750687]
6. Buffler PA, Kwan ML, Reynolds P, Urayama KY. Environmental and genetic risk factors for childhood leukemia: appraising the evidence. Cancer Invest. 2005; 23(1):60–75. [PubMed: 15779869]
7. Treviño LR, Yang W, French D, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet. 2009; 41(9):1001–5. [PubMed: 19684603]
8. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet. 2009; 41(9):1006–10. [PubMed: 19684604]
9. Prasad RB, Hosking FJ, Vijayakrishnan J, et al. Verification of the susceptibility loci on 7p12.2, 10q21.2 and 14q11.2 in precursor B-cell acute lymphoblastic leukemia of childhood. Blood. 2010; 115(9):1765–7. [PubMed: 20042726]
10. Vijayakrishnan J, Sherborne AL, Sawangpanich R, Hongeng S, Houlston RS, Pakakasama S. 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. 2010; 51(10):1870–4. [PubMed: 20919861]
11. Pastorczak A, Górniak P, Sherborne A, et al. 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. 2011; 35(11):1534–6. [PubMed: 21889209]
12. Xu H, Yang W, Perez-Andreu V, et al. Novel susceptibility variants at 10p12.31–12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. J Natl Cancer Inst. 2013. 105(10):733–42. [PubMed: 23512250]
13. Wang Y, Chen J, Li J, et al. Association of three polymorphisms in ARID5B, IKZF1 and CEBPE with the risk of childhood acute lymphoblastic leukemia in a Chinese population. Gene. 2013; 524(2):203–7. [PubMed: 23608171]
14. Lin CY, Li MJ, Chang JG, et al. High-resolution melting analyses for genetic variants in ARID5B and IKZF1 with childhood acute lymphoblastic leukemia susceptibility loci in Taiwan. Blood Cells Mol Dis. 2014; 52(2–3):140–5. [PubMed: 24200646]
15. Migliorini G, Fiege B, Hosking FJ, et al. Variation at 10p12.2 and 10p14 influences risk of childhood B-cell acute lymphoblastic leukemia and phenotype. Blood. 2013; 122(19):3298–307. [PubMed: 23996088]
16. Gharbi H, Ben Hassine I, Soltani I, et al. 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. 2016; 33(3):157–67. [PubMed: 27184773]
17. Bhandari P, Ahmad F, Mandava S, Das BR. 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. 2016; 17(8):3989–95. [PubMed: 27644650]
18. Bhatia S, Sather HN, Heerema NA, Trigg ME, Gaynon PS, Robison LL. Racial and ethnic differences in survival of children with acute lymphoblastic leukemia. Blood. 2002; 100(6):1957–64. [PubMed: 12200352]
19. Kadan-Lottick NS, Ness KK, Bhatia S, Gurney JG. Survival variability by race and ethnicity in childhood acute lymphoblastic leukemia. JAMA. 2003; 290(15):2008–14. [PubMed: 14559954]
20. Chokkalingam AP, Hsu LI, Metayer C, et al. Genetic variants in ARID5B and CEBPE are childhood ALL susceptibility loci in Hispanics. Cancer Causes Control. 2013; 24(10):1789–95. [PubMed: 23836053]
21. Xu H, Cheng C, Devidas M, et al. ARID5B genetic polymorphisms contribute to racial disparities in the incidence and treatment outcome of childhood acute lymphoblastic leukemia. J Clin Oncol. 2012; 30(7):751–7. [PubMed: 22291082]
22. Hsu LI, Chokkalingam AP, Briggs FB, et al. 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. 2015; 26(4):609–19. [PubMed: 25761407]

23. Walsh KM, Chokkalingam AP, Hsu LI, et al. Associations between genome-wide Native American ancestry, known risk alleles and B-cell ALL risk in Hispanic children. Leukemia. 2013; 27(12):2416–9. [PubMed: 23615557]
24. Martinez-Marignac VL, Valladares A, Cameron E, et al. Admixture in Mexico city: implications for admixture mapping of type 2 diabetes genetic risk factors. Hum Genet. 2007; 120(6):807–19. [PubMed: 17066296]
25. Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour. 2009; 9(5):1322–32. [PubMed: 21564903]
26. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21(2):263–5. [PubMed: 15297300]
27. Evans TJ, Milne E, Anderson D, et al. Confirmation of childhood acute lymphoblastic leukemia variants, ARID5B and IKZF1, and interaction with parental environmental exposures. PLoS One. 2014; 9(10):e110255. [PubMed: 25310577]
28. Healy J, Richer C, Bourgey M, Kritikou EA, Sinnett D. Replication analysis confirms the association of ARID5B with childhood B-cell acute lymphoblastic leukemia. Haematologica. 2010; 95(9):1608–11. [PubMed: 20460642]
29. Archer NP, Perez-Andreu V, Stoltze U, et al. Family-based exome-wide association study of childhood acute lymphoblastic leukemia among Hispanics confirms role of ARID5B in susceptibility. PLoS One. 2017; 12(8):e0180488. [PubMed: 28817678]
30. Al-Absi B, Razif MFM, Noor SM, et al. Contributions of IKZF1, DDC, CDKN2A, CEBPE, and LMO1 gene polymorphisms to acute lymphoblastic leukemia in a Yemeni population. Genet Test Mol Biomarkers. 2017; 21(10):592–99. [PubMed: 28768142]
31. Bell JI. Single nucleotide polymorphims and disease gene mapping. Arthritis Res. 2002; 4(suppl 3):S273–8. [PubMed: 12110147]
32. Conomos MP, Laurie CA, Stilp AM, et al. Genetic diversity and association studies in US Hispanic/Latino populations: applications in the Hispanic community health study/study of Latinos. Amer J Hum Genet. 2016; 98(1):165–84. [PubMed: 26748518]
33. Prince AL, Patterson N, Yu F, et al. A genomewide admixture map for Latino populations. Am J Hum Genet. 2007; 80(6):1024–36. [PubMed: 17503322]
34. Linabery AM, Blommer CN, Spector LG, Davies SM, Robison LL, Ross JA. ARID5B and IKZF1 variants, selected demographic factors, and childhood acute lymphoblastic leukemia: a report from the Children's Oncology Group. Leuk Res. 2013; 37(8):936–42. [PubMed: 23692655]

Table 1.-

Clinical characteristics of the patients

Clinical characteristics	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			
$<50\times10^9$/L	228 (80·3)	3 (21·4)	231 (77·5)
$\geq50\times10^9$/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)

Table 2.-

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

SNPs	384 Controls n (%)	298 patients n (%)	*p*	284 B-ALL patients n (%)	*p*	14 T-ALL patients n (%)	*p*
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	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	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	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	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	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	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	0·48	0·55	0·0757	0·55	0·0507	0·43	0·7899

p values by Fisher exact test

Statistcally significant *p* value (bold)

Table 3.-

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

	ALL patients			B-ALL patients		
SNPs	**Risk allele frequency**	**OR (95% CI)**	***p***	**Risk allele frequency**	**OR (95% CI)**	***p***
rs10821936						
T *vs* C	0·68	2·06 (1·64 to 2·57)	**$1·72e^{-10}$**	0·69	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						
G *vs* A	0·71	1·76 (1·40 to 2·22)	**$9·15e^{-07}$**	0·72	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						
T *vs* G	0·68	2·01 (1·61 to 2·52)	**$4·90e^{-10}$**	0·68	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						
G *vs* A	0·37	1·36 (1·08 to 1·71)	**0·0075**	0·38	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						
G *vs* C	0·38	1·23 (0·98 to 1·54)	0·0648	0·40	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						
A *vs* G	0·40	1·32 (1·06 to 1·65)	**0·0128**	0·41	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						
T *vs* C	0·55	1·32 (0·71 to 1·64)	**0·0093**	0·55	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**

p values by Fisher exact test

Statistcally significant *p* value (bold)