

RESEARCH ARTICLE

Editorial Process: Submission:02/22/2023 Acceptance:08/09/2023

Genes of Predisposition to Childhood Beta-Cell Acute Lymphoblastic Leukemia in the Kazakh Population

Gulnara Svyatova¹, Riza Boranbayeva², Galina Berezina^{1*}, Lyazat Manzhuova²,
Alexandra Murtazaliyeva¹

Abstract

Background: Today, acute lymphoblastic leukemia is one of the most common malignant diseases of the hematopoietic system. The genetic predisposition to ALL is not fully explored in various ethnic populations. **Objective:** The study aimed to conduct a comparative analysis of the population frequencies of alleles and genotypes of polymorphic gene variants: immune regulation *GATA3* (*rs3824662*); transcription and differentiation of B cells: *ARID5B* (*rs7089424*, *rs10740055*), *IKZF1* (*rs4132601*); differentiation of hematopoietic cells: *PIP4K2A* (*rs7088318*); apoptosis: *CEBPE* (*rs2239633*), tumor suppressors: *CDKN2A* (*rs3731249*), *TP53* (*rs1042522*); carcinogen metabolism: *CBR3* (*rs1056892*), *CYP1A1* (*rs104894*, *rs4646903*), according to genome-wide association studies analyses associated with the risk of developing pediatric beta-cell acute lymphoblastic leukemia (B-cell ALL), in an ethnically homogeneous population of Kazakhs with studied populations. **Methods:** The genomic database consists of 1800 conditionally healthy persons of Kazakh nationality, genotyped using OmniChip 2.5-8 Illumina chips at the deCODE genetics as part of the InterPregGen 7 project of the European Union (EU) framework program under Grant Agreement No. 282540. **Results:** High population frequencies of single nucleotide polymorphism (SNP) minor alleles identified for immune regulation genes – *GATA3 rs3824662* – 42.5%; transcription and differentiation of B-cells genes – *ARID5B rs7089424* – 33.1% and *rs10740055* – 48.5%, which suggests their significant genetic contribution to the risk of development and prognosis of the effectiveness of B-cell ALL therapy in the Kazakh population. The significantly lower population frequency of the minor allele G *rs1056892 CBR3* gene – 38.6% in the Kazakhs suggests its significant protective effect in reducing the risk of childhood B-cell ALL and the smaller number of cardiac complications after anthracycline therapy. **Conclusion:** The obtained results will serve as a basis for developing effective methods for predicting the risk of development, early diagnosis, and effectiveness of treatment of B-cell ALL in children.

Keywords: B-cell lymphoblastic leukemia in children- SNP- genome-wide association studies analysis- minor alleles

Asian Pac J Cancer Prev, 24 (8), 2653-2666

Introduction

Acute lymphoblastic leukemia (ALL) is a malignant disease of the hematopoiesis system, which manifests by the uncontrolled proliferation of immature lymphoid cells (lymphoblasts). ALL is the most frequent oncological disease of childhood and adolescence in economically developed countries, while B-cell ALL accounts for 80% to 90% of all cases (Belson et al., 2007). In contrast to the sufficiently studied chronic lymphoblastic leukemia (CLL), the genetic predisposition to ALL is not explored fully in various ethnic populations. Several studies have found population differences in the frequency of B-cell ALL, which are probably, relate to environmental factors and genetic background.

Thus, in children of Hispanic origin, the incidence of ALL is 10-30% higher than in non-Hispanic whites and

almost twice as higher in non-Hispanic blacks (Yamamoto and Goodman, 2008; Walsh et al., 2013). At the same time, children of Hispanic origin with ALL have a lower 5-year survival rate and a higher relapse rate than non-Hispanic whites (Kadan-Lottick et al., 2003; Yang et al., 2011). The existing differences in the geographical distribution of the incidence of ALL (Wu et al., 2009) are probably related to the difference in external risk factors, as well as differences in the unique population frequencies of significant polymorphisms associated with ALL. The external adverse effects on the risk of ALL also demonstrated contradictory data (Greaves, 2006).

The leading strategy to detect SNPs associated with B-cell ALL is the GWAS, principles are described in various publications. The first GWAS predisposition to ALL identified risk genes *ARID5B*, *IKZF1* (Papaemmanuil et al., 2009; Trevino et al., 2009), and *CEBPE* (Papaemmanuil

¹Republican Medical Genetic Consultation, Scientific Center of Obstetrics, Gynecology and Perinatology, 050020, 125 Dostyk Ave., Almaty, Kazakhstan. ²Scientific Center of Pediatrics and Pediatric Surgery, 050060, 146 Al-Farabi Ave., Almaty, Kazakhstan. *For Correspondence: g.berezina9459-1@scientific-community.com

et al., 2009) in children of European origin. Following GWAS with larger sample sizes and ethnic diversity found additional significant SNPs in *PIP4K2A*, *GATA3*, *ARID5B*, and *IKZF1* (Xu et al., 2013; Perez-Andreu et al., 2013; Sherborne et al 2010; Migliorini et al 2013) showed high genetic contribution to B-cell ALL regardless of the ethnic origin of patients, *CEBPE* and *CDKN2A* SNPs were significant mainly in European patients with ALL (Xu et al., 2013; Perez-Andreu et al., 2013). Several independent GWAS in ethnically diverse populations identified risk genes – *ARID5B*, *IKZF1*, *CEBPE*, *PIP4K2A*, *CDKN2A*, *GATA3* associated with shorter survival ($p < 0.017$), increased recurrence rate ($p < 0.008$) and poor prognosis ($p < 0.028$) of the course of childhood ALL (Trevino et al., 2009; Xu et al., 2013; Sherborne et al., 2010; Mosad et al., 2016), the significance was confirmed in subsequent replication studies (Vijayakrishnan et al., 2019; Wiemels et al., 2018; Vijayakrishnan et al., 2018; Qian et al., 2019; de Smith et al., 2019).

Conducted genome-wide studies (GWAS) of the search for associations with the risk of ALL have ambiguous results indicating a different etiology of the existing subtypes of ALL with their extreme ethnic variability. 96% of GWAS analyses that conducted in European population is impossible to extrapolate to other ethnic groups. GWAS in non-European and Asian populations are very few and sometimes contradictory, indicating the need for further studies with a large sample size in certain ethnic groups. Authors planned comparative analysis of the population alleles and genotypes frequencies GWAS associated with childhood B-cell ALL, which will allow suggesting significant genetic markers, predict the population frequency, the effectiveness of therapy, the probabilistic outcomes in the Kazakh population. The goals of the study are to identify genetic mutations or variations that increase the risk of developing beta-cell acute lymphoblastic leukemia in children in the Kazakh population. To investigate the frequency and distribution of these genetic variations in the Kazakh population. To explore the possible interactions between genetic factors and environmental factors that may contribute to the development of childhood beta-cell acute lymphoblastic leukemia. To understand the underlying biology of childhood beta-cell acute lymphoblastic leukemia and potential targets for therapy or prevention. To provide insights that can be used to develop personalized treatment plans for patients with childhood beta-cell acute lymphoblastic leukemia based on their genetic profile.

In this study, the authors focused on a comparative analysis of the population alleles and genotypes frequencies of SNPs: immune regulation *GATA3* rs3824662; transcription and differentiation of B-cells: *ARID5B* rs7089424, rs10740055, *IKZF1* rs4132601; differentiation of hematopoietic cells: *PIP4K2A* rs7088318; apoptosis: *CEBPE* rs2239633, tumor suppressors: *CDKN2A* rs3731249, *TP53* rs1042522; metabolism of carcinogens: *CBR3* rs1056892, *CYP1A1* rs104894, rs4646903, according to GWAS risk-associated childhood B-cell ALL, in an ethnically homogeneous Kazakh population with previously studied world populace.

Materials and Methods

The research material is genomic information of 1,800 conditionally healthy persons of Kazakh nationality (1,159 women, 328 men, and 313 children). All the subjects signed informed consent to participate; for the children – one of the parents signed the informed consent. The criteria for selection to the population control group are ethnicity – Kazakhs, including grandparents; the ability of the subject to make an independent decision on consent to participate in the project. Exclusion criteria – in the anamnesis confirmed by medical documents – hypertension, stroke, diabetes mellitus types 1, 2, requiring medical treatment. The deoxyribonucleic acid (DNA) is stored in the JSC “Scientific Center of Obstetrics, Gynecology, and Perinatology” in “Miras” biobank. The biobank “Miras” was created as part of the project “Genetic studies of PE (preeclampsia) in populations of Central Asia and Europe” (InterPregGen) 7 of the framework program of the European Commission Grant Agreement No. 282540.

For the selection of significant SNPs associated with the risk of developing childhood B-cell ALL, we use information about possible candidate genes deposited in public databases: National Center for Biotechnology Information (2021), Ensembl (2021), the Database of Genotypes and Phenotypes (dbGaP) (2021), National Human Genome Research Institute (2021), HapMap (2021), 1000 Genomes Project (2021), GWAS Catalog (2021), Clinical Trials Gov (2021), The integrated Cancer Repository for Cancer Research (iCaRe2) (2021). According to GWAS data we conducted population frequencies comparative analysis of significant SNPs associated with B-cell ALL and selected 11 SNPs of 9 genes. Due to the fact that Illumina Omni Chip 2.5-8 does not contain *GATA3* rs3824662, *CYP1A1* rs104894 and rs4646903, we included these SNPs using PLINK (2021) and Linkage Disequilibrium Calculator (2021) with $LD \geq 0.9$.

DNA was isolated by M-PVA magnetic particle separation on a Prepito automated analyzer (PerkinElmer) to isolate Chemagic Prepito nucleic acids (Wallac, Finland) using the Prepito DNA CytoPure reagent kit. Genotyping for ~2.5 million SNPs was performed in deCODE genetics using Illumina Omni Chip 2.5-8 as part of the InterPregGen project. Genotyping quality control was carried out with the exclusion of SNPs with MAF (minimal allele frequency) below 1%, call rate <98%, significantly less than $p < 5 \times 10^{-8}$, deviation from the Hardy-Weinberg equilibrium ($P < 0.05$) (Manning, 2012; Hirschhorn and Gajdos, 2011; Berndt, 2013). In this article used PLINK software for statistical calculations of allele and genotype frequencies, significance tests, and analysis of the nonparametric criterion χ^2 . The assessment of the correspondence of the obtained genotype frequencies to the Hardy-Weinberg equilibrium law was calculated using the HWE (Hardy-Weinberg equilibrium) test function of the PLINK program (Purcell, 2007). The study was approved by the local Ethics Commission of the Non-Profit JSC “Asfendiyarov Kazakh National Medical University” (Almaty, Kazakhstan), application No. 1189

dated September 28, 2021.

Results

The result of GWAS meta-analysis is particular scientific interest that combined data from four GWAS studies of B-cell ALL: UK GWASI, German GWAS, UK GWASII, and COG_SJGWAS (Vijaykrishnan et al., 2019), based on ~10 million SNPs of 5,321 patients with B-cell ALL and 16,666 controls of European origin. Six newly identified SNPs replicated in an independent sample of 2,237 cases and 3,461 control groups (COG_SJGWAS non-European Americans); SNPs *TLE1 rs76925697*, *IRF1 rs886285*, *BAK1 rs210143*, *IGF2BP1 rs10853104* showed a significant contribution ($p < 0.05$) to the risk developing B-cell ALL (Vijaykrishnan et al., 2019). Immune regulation genes. *GATA3* (GATA Binding Protein 3, *rs3824662*) is an immune regulatory gene encoding the GATA Binding Protein 3, a transcription activator and regulator of the differentiation process of T- and B-cells, natural killer cells, which plays an important role in hematopoietic function (Yamamoto and Goodman, 2008; Walsh et al., 2013; Kadan-Lottick et al., 2003; Yang et al., 2011; Greaves, 2006; Xu et al., 2013; Perez-Andreu et al., 2013).

GWAS study 863370 SNPs in 2,597 children with B-cell ALL from the cohort of the Children's Oncology Group (COG) and 6,661 healthy children found 2 significant SNPs *GATA3: rs3824662* (OR = 3.85, 95% CI: 2.71-5.47, $p = 2.17 \times 10^{-14}$) and *rs3781093* (OR = 3.45, 95% CI: 2.42-4.93, $p = 4.94 \times 10^{-12}$), which were in a strong linkage disequilibrium (LD, $r^2 = 0.94$, $D = 1$) at the same locus (Perez-Andreu et al., 2015; Zhang et al., 2020). *GATA3 rs3824662* minor allele was significantly associated with an increased risk of B-cell ALL relapses and a poor survival to justify treatment intensification

(Zhang et al., 2020).

Similar result are presented in GWAS of European and Latin American pediatric patients with B-cell ALL (Papaemmanuil et al., 2009; Trevino et al., 2009; Migliorini et al., 2013), which confirmed a significant genetic contribution *GATA3 rs3824662* to the development and risk of relapses in children with ALL in both ethnic groups (Walsh et al., 2013; Xu et al., 2021; Moriyama et al., 2018). Carriage of minor allele *rs3824662* explains the 1.11-fold increasing the risk of B-cell ALL and the more severe disease prognosis in Latin Americans compared to European patients (Migliorini et al., 2013). In 2 recent GWAS of Hispanic Latin children with B-cell ALL frequency of risk genotypes *rs3824662 GATA3* was 30% higher in Latin Americans compared to Europeans (Xu et al., 2021), which caused an almost 4-fold increase risk of leukemia. The studies confirm the presence of ethnic differentiation in carriage of risk alleles of significant SNPs of the *ARID5B*, *PIP4K2A* and *GATA3* genes in ALL, which possibly causes racial differences in the incidence of ALL in African, European and Latin American populations (Vijaykrishnan et al., 2018; Moriyama et al., 2015). B-cell transcription and differentiation genes. The *ARID5B* gene (AT-Rich Interaction Domain 5B) encodes protein forms a histone H3K9Me2 demethylase complex with PHD finger protein 2, as a transcription regulator, participates in the growth and differentiation of B-lymphocyte precursors (Wilsker et al., 2002).

Conducted GWAS of 907 patients with ALL and 2398 control group identified loci of ALL risk: *IKZF1 rs4132601* (7p12.2), *ARIDB5 rs7089424* (10q21.2), and *CEBPE rs2239633* (14q11.2), all three significant SNPs correspond to genes that regulate transcription and differentiation of B-cell precursors (Papaemmanuil et al., 2009). The association of ALL with the detected loci was subsequently confirmed by a multi-ethnic

Table 1. Characteristics of 11 SNPs 9 Genes GWAS Associated with Childhood B-cell ALL

No.	Gene name	Chromosome	rs	Position
Immune regulation gene				
1	<i>GATA3*</i> GATA Binding Protein 3	10	rs3824662	8062245
B-cell transcription and differentiation genes				
2	<i>ARID5B</i> AT-Rich Interaction Domain 5B	10	rs7089424	61992400
3	<i>ARID5B</i>	10	rs10740055	61958720
4	<i>IKZF1</i> IKAROS Family Zinc Finger 1	7	rs4132601	50402906
Differentiation of hematopoietic cells gene				
5	<i>PIP4K2A</i> Phosphatidylinositol-5-Phosphate 4-Kinase Type 2 Alpha	10	rs7088318	22564019
Apoptosis gene				
6	<i>CEBPE</i> CCAAT Enhancer Binding Protein Epsilon	14	rs2239633	23119848
Tumor suppressor genes				
7	<i>CDKN2A</i> Cyclin Dependent Kinase Inhibitor 2A	9	rs3731249	21970917
8	<i>TP53</i> Tumor Protein p53	17	rs1042522	7676154
Carcinogen metabolism genes				
9	<i>CBR3</i> Carbonyl Reductase 3	21	rs1056892	36146408
10	<i>CYP1A1*</i> Cytochrome P450 Family 1 Subfamily A Member 1	15	rs1048943	74720644
11	<i>CYP1A1*</i>	15	rs4646903	74719300

Note: rs is the SNP identifier; * is not included in IlluminaOmni 2.5-8 Chip.

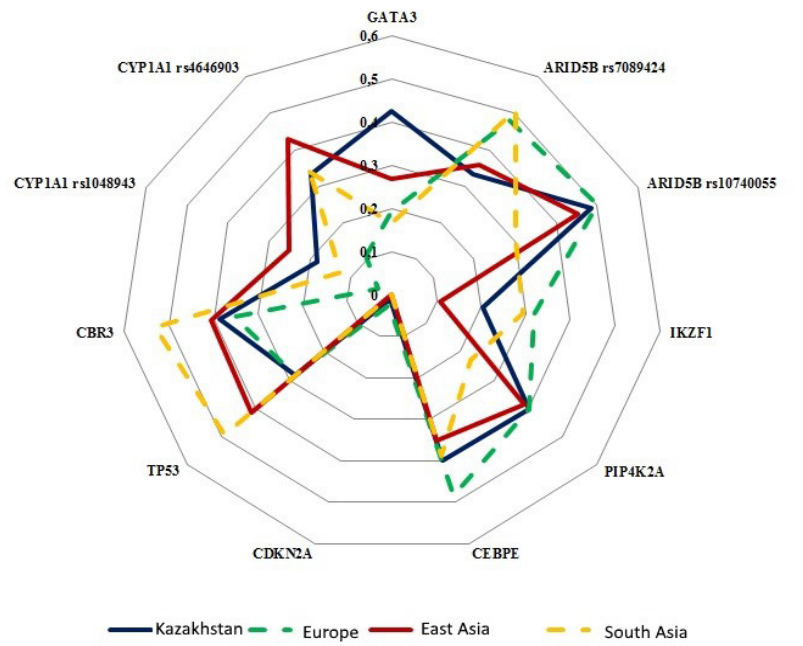


Figure 1. Comparative Analysis of Population Minor Alleles Frequencies of 11 SNPs GWAS associated with B-cell ALL

GWAS (Xu et al., 2013) of 1,605 children with ALL and 6,661 control individuals, followed by replicative genotyping of 845 patients and 4,316 controls in European Americans, African Americans, and Hispanic Americans ($P = 0.001, 0.009, \text{ and } 0.04$, respectively). *ARID5B* rs7089424 and rs10740055 were found in previous GWAS of ALL (Papaemmanuil et al., 2009; Trevino et al., 2009; Ellinghaus et al., 2012) and demonstrated a statistically significant association with ALL ($P < 5.0 \times 10^{-8}$)

(Papaemmanuil et al., 2009; Trevino et al., 2009; Orsi et al., 2012), the strongest in patients of the white and Latin population of America (Xu et al., 2013).

The *IKZF1* gene (IKAROS Family Zinc Finger 1, rs4132601) (IKAROS Family Zinc Finger 1) encodes the IKAROS protein transcription factor, expressed in hematopoietic stem cells that regulate proliferation and differentiation of B-cells during hematopoiesis (Kano et al., 2008). It is well-known that the unfavorable

Table 2. Frequencies of Alleles and Genotypes of the Studied SNPs GWAS associated with Childhood B-cell ALL in the Kazakh Population

Gene name	rs	MAF	N	A1	A2	GENO
Immune regulation gene						
<i>GATA3</i> *	rs3824662	0.4255	1,799	A	C	589/889/321
B-cell transcription and differentiation genes						
<i>ARID5B</i>	rs7089424	0.3312	1,801	G	T	813/783/205
<i>ARID5B</i>	rs10740055	0.485	1,798	A	C	481/890/427
<i>IKZF1</i>	rs4132601	0.2029	1,801	G	T	1147/577/77
Differentiation of hematopoietic cells gene						
<i>PIP4K2A</i>	rs7088318	0.4037	1,796	C	A	648/846/302
Apoptosis gene						
<i>CEBPE</i>	rs2239633	0.3997	1,800	A	G	642/877/281
Tumor suppressor genes						
<i>CDKN2A</i>	rs3731249	0.009733	1,798	T	C	1763/35/0
<i>TP53</i>	rs1042522	0.2832	1,790	C	G	919/728/143
Carcinogen metabolism genes						
<i>CBR3</i>	rs1056892	0.3864	1,800	A	G	694/821/285
<i>CYP1A1</i> *	rs1048943	0.1824	1,782	C	T	1198/518/66
<i>CYP1A1</i> *	rs4646903	0.3325	1,800	G	A	825/753/222

Note: rs is the SNP identifier; MAF is the frequency of the minor allele; N is the number of genotyped individuals; A1 is the common allele and A2 is the minor allele; GENO is the number of identified genotypes; * is not included in IlluminaOmni 2.5-8 Chip.

Table 3. Genotype Distribution to the Hardy-Weinberg Equilibrium of 11 SNPs GWAS Associated with Childhood B-cell ALL in the Kazakh Population

Gene name	rs	N	GENO	O(HET)	E(HET)	P
Immune regulation gene						
<i>GATA3</i>	rs3824662	1,799	589/889/321	0.4942	0.4889	0.6645
B-cell transcription and differentiation genes						
<i>ARID5B</i>	rs7089424	1,801	813/783/205	0.4348	0.443	0.4253
<i>ARID5B</i>	rs10740055	1,798	481/890/427	0.495	0.4995	0.7057
<i>IKZF1</i>	rs4132601	1,801	1147/577/77	0.3204	0.3235	0.6626
Differentiation of hematopoietic cells gene						
<i>PIP4K2A</i>	rs7088318	1,796	648/846/302	0.471	0.4814	0.3776
Apoptosis gene						
<i>CEBPE</i>	rs2239633	1,800	642/877/281	0.4872	0.4799	0.5556
Tumor suppressor genes						
<i>CDKN2A</i>	rs3731249	1,798	1763/35/0	0.01947	0.01928	1
<i>TP53</i>	rs1042522	1,790	919/728/143	0.4067	0.406	1
Carcinogen metabolism genes						
<i>CBR3</i>	rs1056892	1,800	694/821/285	0.4561	0.4742	0.1116
<i>CYP1A1*2C</i>	rs1048943	1,782	1198/518/66	0.2907	0.2982	0.3014
<i>CYP1A1</i>	rs4646903	1,800	825/753/222	0.4208	0.4461	0.0958

Note: rs is the SNP identifier; MAF is the frequency of the minor allele; N is the number of genotyped individuals; GENO is the number of identified genotypes; O (HET) is the expected heterozygosity according to the Hardy-Weinberg equilibrium; E (HET) is the observed heterozygosity according to the Hardy-Weinberg equilibrium; P is p-value; * is not included in IlluminaOmni 2.5-8 Chip.

polymorphism *IKZF1 rs4132601* is associated with the development of childhood B-cell ALL 9 (Sellars et al., 2011). It assumed that the G allele reduces stability and expression level of mRNA, which delays the B-cells maturation and increases ALL risk (Górniak et al., 2014). A lower expression of *IKZF1* in the presence of GG rs4132601 may contribute to a more aggressive clinical course of the disease (Lautner-Csorba et al., 2012) and negative effect on B-cell ALL (Song et al., 2015).

Independent replicative genotyping of *IKZF1 rs4132601*, *rs111978267* of 170 children with B-cell ALL and 150 control subjects in the Tunisian population revealed a significant association of G allele ($p = 0.029$) with an increased 1.54-fold risk of B-cell ALL. The association of rs4132601 with B-cell ALL was observed within the codominant ($p = 0.009$), recessive ($p = 0.006$) and additive ($p = 0.027$) genetic models. On the contrary, there is no significant association with the risk of developing childhood ALL for *rs111978267* (Mahjoub et al., 2019). Multiethnic GWAS meta-analysis of 15 case-control studies of 8,333 patients with B-cell ALL and 36036 control cases gave contradictory results (Li et al., 2015). Significant associations between unfavorable genotypes carriage of *IKZF1 rs4132601* and the development of B-cell ALL found for Europeans and Latin Americans (Orsi et al., 2012; Vijayakrishnan et al., 2010); were not confirmed among Asian and African populations (Lin et al., 2013); for Indian children play a protective role in reducing the risk of developing B-cell ALL (Xu et al., 2013; Perez- Rudant et al., 2015). Thus, conducted meta-GWAS provides evidence that the *IKZF1 rs4132601* predisposes to the occurrence of B-cell ALL, especially in European populations. The studies for B-cell

ALL associations with the carrier of this polymorphism in non-Europeans and Asian are insufficient and sometimes contradictory, which indicates the need for further replicative studies with a large sample size in various ethnic groups.

Differentiation of hematopoietic cells gene. The *PIP4K2A* gene (Phosphatidyl inositol-5-Phosphate 4-Kinase Type 2 Alpha, *rs7088318*) encodes the same name protein that participates in hematopoiesis by regulating secretion, proliferation, and differentiation of hematopoietic cells (Fiume et al., 2015). Conducted GWAS has shown the genetic contribution of *PIP4K2A rs7088318* to the risk of ALL in children in ethnically diverse populations (Xu et al., 2013; Shin et al., 2019; Zhang et al., 2019). One of the first multiethnic GWAS of 1605 children with ALL and 6661 control subjects found a new significant locus *PIP4K2A rs7088318* ($p = 1.1 \times 10^{-10}$). The identified association, was confirmed in three replication cohorts of European Americans, African Americans, and Latin Americans ($p = 0.001$, 0.009, and 0.04, respectively) (Xu et al., 2013). The opposite results are obtained after replicative genotyping this polymorphism in 191 patients with B-cell ALL and 342 healthy individuals in the Spanish population. There is no statistically significant relationship between *rs7088318* and the risk of B-cell ALL in any of the analyzed genetics models (Lopez-Lopez et al., 2013). It is assumed, that the discrepancies in the results are related to differences in the population frequencies of risk alleles of significant polymorphisms associated with ALL. Thus, the frequency of minor alleles in the studied control population was significantly higher (MAF = 0.64) than in control European-American population (MAF = 0.59)

Table 4. Comparative Analysis of SNPs Allelic Frequencies GWAS Associated with Childhood B-cell ALL in World Populations

Population	N	MAF	χ^2	P
Immune regulation gene				
<i>GATA3 rs3824662</i>				
Kazakhstan	1799	0.4255		
Europe*	503	0.193	91.024	<0.001
East Asia*	504	0.269	40.187	<0.001
South Asia*	489	0.166	111.606	<0.001
B-cell transcription and differentiation genes				
<i>ARID5B rs7089424</i>				
Kazakhstan	1801	0.3312		
Europe*	503	0.329	0.004	0.951
East Asia*	504	0.359	1.346	0.246
South Asia*	489	0.509	52.143	<0.001
<i>ARID5B rs10740055</i>				
Kazakhstan	1798	0.485		
Europe*	503	0.501	0.403	0.526
East Asia*	504	0.454	1.479	0.224
South Asia*	489	0.301	52.899	<0.001
<i>IKZF1 rs4132601</i>				
Kazakhstan	1801	0.2029		
Europe*	503	0.318	29.774	<0.001
East Asia*	504	0.108	24.005	<0.001
South Asia*	489	0.293	17.953	<0.001
Differentiation of hematopoietic cells gene				
<i>PIP4K2A rs7088318</i>				
Kazakhstan	1796	0.4037		
Europe*	503	0.404	0.000	0.997
East Asia*	504	0.388	0.358	0.550
South Asia*	489	0.230	50.498	<0.001
Apoptosis gene				
<i>CEBPE rs2239633</i>				
Kazakhstan	1800	0.3997		
Europe*	503	0.483	11.312	<0.001
East Asia*	504	0.35	4.183	0.041*
South Asia*	489	0.389	0.191	0.663
Tumor suppressor genes				
<i>CDKN2A rs3731249</i>				
Kazakhstan	1798	0.009733		
Europe*	503	0.025	7.414	1
East Asia*	504	0	5.085	0.025*
South Asia*	489	0	4.934	0.027*
<i>TP53 rs1042522</i>				
Kazakhstan	1790	0.2832		
Europe*	503	0.285	0.002	0.964
East Asia*	504	0.414	157.823	<0.001
South Asia*	489	0.492	76.528	<0.001
Carcinogen metabolism genes				
<i>CBR3 rs1056892</i>				
Kazakhstan	1800	0.3864		

Table 4. Continued

Population	N	MAF	χ^2	P
<i>CBR3 rs1056892</i>				
Europe*	503	0.354	1.795	0.181
East Asia*	504	0.405	0.542	0.462
South Asia*	489	0.532	33.255	<0.001
<i>CYP11A1*2C rs1048943</i>				
Kazakhstan	1782	0.1824		
Europe*	503	0.035	134.960	<0.001
East Asia*	504	0.252	1.740	0.188
South Asia*	489	0.127	49.331	<0.001
<i>CYP11A1 rs4646903</i>				
Kazakhstan	1800	0.3325		
Europe*	503	0.107	98.340	<0.001
East Asia*	504	0.430	16.458	<0.001
South Asia*	489	0.339	0.077	0.781

Note: χ^2 , chi-square; *, Ensembl genome database project (1000 Genomes Project, 2021).

(Xu et al., 2013), which indicates that *PP4K2A rs7088318* is associated with the risk of B-cell ALL in children in several ethnic groups, but is not a diagnostically significant genetic marker in the Spanish population (Lopez-Lopez et al., 2013).

Apoptosis gene. The *CEBPE* protein encoded by the *CEBPE* gene (CCAAT Enhancer Binding Protein Epsilon rs2239633) belongs to the bZIP transcription factor in terminal differentiation, functional maturation of myeloid precursor cells committed to neutrophils, monocytes, and granulocytes (Akasaka et al., 2007). A growing number of published studies have devoted the relationship between the carriage *CEBPE rs2239633* and ALL risk (Trevino et al., 2009; Ellinghaus et al., 2012; Kano et al., 2008; Prasad et al., 2010; Wang et al., 2013), but the results were equivocal and sometimes contradictory. *CEBPE rs2239633* is significantly associated with the risk of B-cell ALL developing in children and adults with OR = 1.1-1.6 (Papaemmanuil et al., 2009; Gharbi et al., 2016). The first GWAS conducted by L.R. Trevino et al. (2009) identified eighteen significant SNPs in 307944 patients with B-cell ALL, including *CEBPE rs2239633*, which increased the risk of B-cell ALL by 1.5-1.6 times. There are also opposite results. Wang et al. (2013) conducted a GWAS meta-analysis of 11 case-control studies, including 5639 patients with ALL and 10036 control subjects, but did not find this association. At the same time, the analysis of ethnic stratification showed a significant association of *CEBPE rs2239633* with childhood ALL in the European and Latin American groups but not confirmed for children with ALL in the Asian sample. It found a decreased risk of B-cell ALL in children in the European population with the minor allele *CEBPE rs2239633* carrier, which was probably conditioned by the difference in sample size and ethnic composition examined. J. Sun et al. (2015) held a meta-analysis of 22 published studies involving 6,152 children with B-cell ALL and 11,739 healthy children. It found a decreased risk of B-cell ALL in children in

the European population with the minor allele *CEBPE rs2239633* carrier, probably due to the difference in sample size and examined ethnic composition.

In light of the conflicting results, an updated meta-analysis based on publications in Web of Science, PubMed, Cochrane Library, Embase, and China National Knowledge Infrastructure (CNKI) included 20 studies with a total sample of 7014 patients with B-cell ALL and 16,428 control subjects (Liu et al., 2021). A significant association was found between the carriage of *CEBPE rs2239633* and the B-cell ALL risk in children (OR = 1.19, 95% CI: 1.11-1.28, $p < 0.01$). The analysis of ethnic stratification also showed a relationship between this SNP and B-cell ALL in childhood in European (OR = 1.19, 95% CI: 1.09-1.30, $p < 0.01$) and Spanish-speaking (OR = 1.39, 95% CI: 1.18-1.63, $p < 0.01$) samples. No significant association was observed in the Asian subgroup (OR = 1.05, 95% CI: 0.90-1.22, $p = 0.53$) (Nakajima et al., 2006). Despite the existing controversies, most studies confirm the presence of ethnic differences in ALL development and show that *CEBPE rs2239633* is associated with B-cell ALL increased risk in the European population.

Tumor suppressor genes. The *CDKN2A* gene (Cyclin-Dependent Kinase Inhibitor 2A *rs3731249*) encodes several proteins, such as p16 (INK4A) and p14 (ARF). Both proteins function as tumor suppressors – they control the growth and intensive cell division and are involved in stopping the division of old cells. Protein p14 (ARF) protects the tumor-suppressive protein p53 from destruction, thereby helping to prevent the formation of tumors (*CDKN2A* cyclin dependent..., 2021). The *CDKN2A rs3731249*, *rs113650570*, and *rs36228834* are in a nonequilibrium coupling (LD = 0.98-1.0) and determines one risk haplotype: minor allele carriage leads to the loss of ability to suppress leukemic transformation and the ALL develop risk (Xu et al., 2015).

Multiethnic genotyping of 321 Latin American children with B-cell ALL and 454 Latin American control subjects from the California Pediatric Leukemia Study (CCLS) was carried out to search for associations of childhood ALL and *CDKN2A* gene polymorphisms, followed by replicative genotyping of an independent group of 980 children of European origin with B-cell ALL and 2624 healthy children from the Pediatric Oncological Group (COG) and the Wellcome Trust Case Management Consortium (WTCC) (Speedy et al., 2014; Walsh et al., 2015). In the CCLS and the COG group, the minor allele C *CDKN2A rs3731249* was associated with an almost 3-fold increase in the B-cell ALL risk (OR = 2.77; 95% CI: 1.58-4.85; $p = 3.78 \times 10^{-4}$) and (OR = 2.99; 95% CI: 2.10⁻⁴.26; $p = 1.51 \times 10^{-9}$) respectively. TaqMan replicative genotyping of the third group of “case-control” children of Latin American and African-American origin (163 cases of ALL, 201 controls) also confirmed 3-fold increase in the risk associated with the carrier *CDKN2A rs3731249* (OR = 3.59; 95% CI: 1.22-10.59; $p = 8.8 \times 10^{-3}$). The meta-analysis of the obtained associations between *rs3731249* and ALL risk achieved genome-wide statistical significance (OR meta = 2.97; 95% CI: 2.22-3.96, $p_{meta} = 1.69 \times 10^{-13}$) (Walsh et al., 2015).

At the same time, similar studies of Arabic-speaking

Tunisians in North Africa showed the absence of an association of *CDKN2A rs3731217* and *CDKN2B rs662463* with ALL in genetic models (codominant, dominant, or recessive) (Gharbi et al., 2016; Mahjoub et al., 2018). Studies by Ch.G. Mullighan et al., (2009) and Lopez-Lopez et al., (2013) have shown that polymorphisms in the *CDKN2A* are associated with an increased risk of leukemia in one ethnic group but not for another. GWAS studies of B-cell ALL in various populations revealed a number of significant loci, while only *CDKN2A rs3731249* achieved genome-wide significance (OR = 2.42, $p = 3.45 \times 10^{-19}$) (Sherborne et al., 2010; Perez-Andreu et al., 2015; Xu et al., 2015). Meta-analysis of GWAS - 24 independent studies of various ethnic groups with a total sample of 7,922 ALL cases and 21,503 control subjects, followed by replicative independent genotyping of 6,295 ALL cases and 24,181 controls, showed a significant association of *CDKN2A rs3731249* with a 2.26-fold increase in ALL risk. The OR in the carriage of the *rs3731217* risk allele varies for different ages and ethnic groups due to the ethnic conditionality of ALL-associated polymorphisms and differences in LD patterns in populations (Orsi et al., 2012; Vijaykrishnan et al., 2010; Xu et al., 2015; Gutierrez-Camino et al., 2017).

The tumor suppressor gene *TP53* (Tumor Protein p53, rs1042522) encodes the protein TP53, which is a negative regulator of cell proliferation and a positive regulator of apoptosis in response to DNA damage. A well-known characteristic of the *TP53* gene is its ability to restrain cell proliferation/differentiation associated with aberrant and uncontrolled oncogene expression. TP53 inactivation induced by unfavorable polymorphism or deletion of the gene contributes to increased oncogene activity and unchecked cell growth (Bieging et al., 2014). Do et al., (2009) studied 114 patients with ALL and 414 control group newborns from Wales (Great Britain) and found a significant contribution of the carriage of G/G genotype TP53 rs1042522 to the development of ALL (OR = 2.9, 95% CI: 1.5-5.6; $p = 0.002$). The results confirmed the possibility of using this polymorphism as a marker of predisposition to ALL development. In the study conducted by Kampouraki et al., (2021) in 86 children with B-cell ALL and 125 control subjects of the European population. The possible association of *CXCL12 rs1801157*, *TP53 rs1042522*, and *CYP1A1*2C rs1048943* with increased susceptibility to the development of B-cell ALL found a higher frequency of heterozygote carriage *CYP1A1*2C* and rare G/G homozygotes of TP53 in children with ALL, no significant differences for the SNP CXCL12.

Carcinogen metabolism genes. Cytochromes P450 (CYP) are a superfamily of enzymes containing heme as a cofactor that functions as monooxygenases. Gene polymorphisms have been identified as risk factors for the development of leukemia in children and are associated with the xenobiotic system, cellular regulation, and the DNA repair system (Berka et al., 2011; Brisson et al., 2015). The *CYP1A1* gene (Cytochrome P450 Family 1 Subfamily A Member 1) encodes the enzyme *CYP1A1* of cytochrome P450, which is investigated actively for

its ability to deactivate compounds with carcinogenic properties. It was shown that *CYP1A1 rs4646903*, *rs1048943* lead to a change in the level of gene expression and affect the stability of RNA messenger, as a result, the enzyme exhibits increased activity, apparently having a decisive effect on the risk of leukemia (Nebert et al., 1996).

In the study of replicative genotyping of the detoxification system – *CYP1A1*, *CYP2D6*, *GSTM1*, and *GSTT1* in 177 children with B-cell ALL and 304 control groups in the French Quebec population, significant associations of the null genotype of *GSTM1* and *CYP1A1 *2C rs1048943* were found, which proved to be significant predictors of the risk of childhood B-cell ALL (OR = 1.8). These results suggest the ALL risk may be related to the metabolism of xenobiotics and, therefore, to adverse environmental influences (Krajinovic et al., 1999). Similar results were obtained in earlier population-based studies (Agha et al., 2013). The effect of *CYP1A1 rs1048943* on the risk of ALL was studied in a meta-analysis of eight multiethnic studies involving 1,734 patients with ALL and 2,194 controls. It was shown that the risk of leukemia was higher in Europeans in the recessive model (OR = 2.23, 95% CI: 1.36–3.68, P = 0.002) but not in the dominant model (OR = 1.22, 95 % CI: 0.95). -1.58, P = 0.12) (Lu et al., 2015). These results are consistent with the findings published by Swinney et al., (2011) conducted a study of Caucasian, Hispanic origin, and African American children with ALL. The authors found that *CYP1A1 rs1048943* was associated with an increased risk of ALL, but after stratification by ethnicity, the risk was statistically significant only for Latin American origin children. However, there are also opposite results. Genotyping pediatric patients with ALL in a genetically homogeneous Kashmir population (Nazki et al., 2012) did not identify a reliable association of *CYP1A1 rs4646903* with the ALL risk, which may be due to the small sample size and the design of the study. In a meta-analysis of 16 studies, which included data on 2,299 patients with ALL and 3209 control cases, a reliable association of *CYP1A1 rs4646903* with the risk of B-cell ALL was confirmed by the dominant model (OR = 1.39, 95% CI: 1.10-1.76, p = 0.006), but not the recessive model (OR = 1.43, 95% CI: 0.76-2.70, p = 0.27) (Zhuo et al., 2012).

It was found that *CYP1A1* is associated with an increased genetic predisposition (Krajinovic et al., 1999) and a worse prognosis (Krajinovic et al., 2002) for childhood B-cell ALL among French-Canadians. Another study showed that the *CYP1A1* significantly increases susceptibility to ALL among Indian children, while the homozygous genotype *rs4646903* gives a 6-fold increase risk and *rs1048943* 4-fold risk of ALL (Joseph et al., 2004). However, the possible role of *CYP1A1* in ALL has not been confirmed in Turkish (Homo sapiens GATA..., 2021), Iranian (Razmkhah et al., 2011), Brazilian (Homo sapiens GATA binding protein..., 2021), and Chinese (Chen et al., 2008) studies. The divergencies obtained in the results of studies on the contribution of *CYP1A1 rs1048943* and *rs4646903* to the risk of ALL developing indicate the existing genetic differences between populations, including the use of different classifications of leukemia: French, American,

and British immunophenotyping systems (Bennett et al., 1976) and the World Health Organisation (Harris et al., 1999). Differences between classification systems can lead to tumor types mixing, inaccurate combined estimates, and cause heterogeneous results. Conducted studies indicate a positive relationship between *CYP1A1* and the ALL risk in childhood. Although the specific causes of childhood leukemia are still not proven, a large number of existing studies suggest that the occurrence of ALL in children is the result of a combination of genetic and environmental factors. Therefore, the study of the relationship between the xenobiotic detoxification system *CYP1A1* and childhood ALL will allow us to determine the possible influence of environmental factors on the risk of ALL.

Table 1 presents the genetic characteristics of 11 SNPs 9 genes selected from the GWAS databases and influence the risk of developing childhood B-cell ALL. ID and location are specified by the SNP on the chromosome. The authors used genomic information of 1800 conditionally healthy Kazakhs from the genomic database of the “Miras” DNA biobank.

Table 2 shows the frequencies of minor alleles of the studied SNPs, GWAS associated with B-cell ALL in children in the Kazakh population.

As can be seen from Table 2, the frequencies of minor alleles of the presented SNPs in the Kazakh sample in the study were: immune regulation *GATA3 rs3824662* – 42.6%; B-cell transcription and differentiation – *ARID5B rs7089424* – 33.1%, *rs10740055* – 48.5%; *IKZF1 rs4132601* – 20.3%; hematopoietic cell differentiation – *PIP4K2A rs7088318* – 40.4%; apoptosis – *CEBPE rs2239633* – 39.9%; tumor suppressors – *CDKN2A rs3731249* – 1.0%; *TP53 rs1042522* – 28.3%; carcinogen metabolism – *CBR3 rs1056892* – 38.6%; *CYP1A1 rs104894* – 18.2%; *CYP1A1 rs4646903* – 33.3%. The lowest population frequency of the minor allele – 1.0% was found for the *CDKN2A rs3731249*, the highest MAF was found for the *ARID5B rs10740055* – 48.5% in the Kazakh population.

Table 3 presented the correspondence of the genotype distribution to the Hardy-Weinberg equilibrium for 11 SNPs associated with the risk of B-cell ALL in the Kazakhs and obtained the result of statistical processing in the PLINK program using the HWE test function. Table 4 shows the results of a comparative analysis of 11 SNP minor alleles frequencies predisposed to the development of B-cell ALL in the Kazakhs compared with previously studied world populations. The frequencies of alleles in the world populations are present according to the 1,000 genomes database (2021), gnomAD (2021), and according publications from world databases, meta-GWAS. It is noteworthy that the sample size – 1800 Kazakhs, turned out to be the highest, which indicates the reliability of the results and the possibility of their extrapolation to the entire Kazakh population.

As shown in Table 4, the frequency of the *GATA3 rs3824662* minor allele in the studied Kazakh sample size was 4.5%, which differed significantly in the populations of Europe – 19.3%, East Asia – 26.9%, South Asia – 16.6% (p<0.001).

Genes for transcription and B-cell differentiation play

a key role in predisposition to B-cell ALL in ethnically diverse populations. Table 4 materials indicate that the allele T *ARID5B* rs7089424 occurs in the Kazakh population with a frequency of 33.1%, which does not significantly differ from the similar indicator in Europe – 32.9% and East Asia – 35.9% ($p > 0.05$). It found significant differences with a higher population frequency of the minor allele in South Asia – 50.9% ($p < 0.001$). Table 4 shows that the carrier of the minor allele C rs10740055 at *ARID5B* in the Kazakh population is 48.5%, which practically had no differences from similar indicators of the European – 50.1% and East Asian – 45.4% populations ($p > 0.05$). The population frequency of the rare allele rs10740055 SNP in the South Asian population was significantly lower – 30.1% ($p < 0.001$). The multiethnic GWAS confirmed the possibility of using *ARID5B* rs7089424 and rs10740055 as reliable prognostic markers of the risk of developing childhood B-cell ALL (Papaemmanuil et al., 2009). In Table 4, the frequency of carrying the minor allele T rs4132601 at *IKZF1* in Kazakhs is 20.3% and occupies an intermediate position compared with previously studied world populations. It significantly exceeds frequency in East Asia – 10.8%, but was lower than the population frequency of T allele Europe – 31.8% and South Asia – 29.3% ($p < 0.001$).

Table 4 shows the population differences in the frequency of the minor allele A rs7088318 at *PIP4K2A*. In Kazakhs, frequency is 40.4%, which has no significant difference with Europeans – 40.4% and East Asians – 38.8% ($p > 0.05$). The lowest frequency of carrying the allele A rs7088318 *PIP4K2A* in the South Asian population – 23.0% ($p < 0.001$). It is well known that the gene *PIP4K2A* is involved in hematopoiesis, thrombopoiesis, terminal maturation of megakaryocytes, and regulation of their size. GWAS of *PIP4K2A* rs7088318 association with the risk of developing B-cell ALL demonstrated a clear dependence on the patients' ethnicity. It confirmed the reliable association in cohorts – Americans of European and Latin American origin, African Americans, but not Hispanic Americans (Xu et al., 2013). Allele frequency *PIP4K2A* rs7088318 is less typical for Hispanic Americans. The genetic heterogeneity of the population reflects the genetic basis of ethnic differences in multifactorial diseases, including B-cell ALL (Linabery and Ross, 2008). Ethnic stratification of GWAS results suggests the most significant association of this polymorphism with B-ALL in European – 40.4%, Kazakhs – 40.4%, and East Asia – 38.8%. This genetic marker is not associated with B-cell ALL in the South Asian population – 23.0%. Consequently, *PIP4K2A* rs7088318 does not seem to be the same general marker of B-cell ALL risk as other previously proposed SNPs of *ARID5B* or *IKZF1*, which was genotyped in many populations and showed its significance.

As shown in Table 4, the frequency of the minor allele G *CEBPE* rs2239633 in the studied sample size of Kazakhs was 39.9%, had no significant differences from the previously studied populations of East – 35.0% and South Asia – 38.9% ($p > 0.05$). Significant differences are obtained for a higher frequency of the unfavorable allele G in the studied populations of Europe – 48.3% ($p < 0.001$).

The *CEBPE* protein plays key role in the terminal differentiation and maturation of myeloid progenitor cells (Nakajima et al., 2006); SNP in the *CEBPE* gene disrupt apoptosis processes, leading to the development of B-cell ALL in children and adults (Papaemmanuil et al., 2009).

The results of numerous GWAS were opposed and convincingly demonstrated ethnic stratification. A significant association was found for *CEBPE* rs2239633 with ALL development in childhood for the European and Latin American patients but not confirmed for children with B-cell ALL in the Asian population (Papaemmanuil et al., 2009; Sun et al., 2015). The possible genetic contribution of this polymorphism in the European patients is associated with a significantly higher population frequency of carrying the minor allele G – 48.3%. Accordingly, *CEBPE* A/G and G/G genotypes are more common in European, while Asian populations demonstrate homogeneous population frequencies. Thus, as shown in Table 3, the population frequency of genotype A/G in Kazakhs was 48.7%, homozygous G/G – 15.6%, which implies a low genetic contribution of *CEBPE* rs2239633 to the risk of childhood B-cell ALL in the Kazakh population. According to the materials, the frequency of the minor allele *CDKN2A* rs3731249 tumor suppressor gene in the Kazakh population was 1.0%, which turned out to be significantly lower than the similar frequency in the European population – 2.5% ($p < 0.05$), but exceeded similar frequencies in East and South Asia, where the risk allele C not detected – 0.0% ($p < 0.05$).

CDKN2A protein functions as a tumor suppressor and inactive excessive uncontrolled cell growth (Mahjoub et al., 2018). Most conducted GWAS have shown that the carriage of the minor allele *CDKN2A* rs3731249 leads to a violation of the ability to suppress leukemic transformation and increases the risk of B-cell ALL in children of European, Latin American, and Spanish origin (Xu et al., 2015; Walsh et al., 2015). The carriage of unfavorable genotypes of C/T and C/C *CDKN2A* rs3731249 increase the risk of B-cell ALL by 2.6-3-fold ($P = 0.002$) (Prasad et al., 2010). There are also opposite results of genomic studies, which did not find statistically significant differences in the frequency of *CDKN2A* rs3731249 in children with B-cell ALL and the control group (Kreile et al., 2016). Large-scale population studies have shown that rs3731249 is an ethnically specific variant, the risk allele C was not detected in Asians (MAF = 0, N = 4254), rarely observed in Africans (MAF = 0.42%, N = 4879), more common in Latin America (MAF = 1.41%, N = 5739) and among Europeans (MAF = 3.52%, N = 32549) according to the largest database ExAC Browser (2021). The minor allele *CDKN2A* rs3731249 increases ALL risk by 3-fold in three independent and ethnically diverse “case-control” samples (Walsh et al., 2015). Consequently, the population frequency of *CDKN2A* rs3731249 in the Kazakh population is 1.0%, significantly exceeding its zero frequency in Asians, indicates a possible high association of this SNP with the risk of development and severity of B-cell ALL in Kazakh children.

As shown in Table 4, the frequency of the minor allele G rs1042522 at TP53 in the Kazakhs was 28.3%, which had no significant differences with the populations of

Europe – 28.5% ($p > 0.05$), but was significantly lower than its frequency of East – 41.4% and South Asians – 49.2% ($p < 0.001$). Tumor suppressor gene TP53 regulates DNA repair and activates apoptosis and gene transcription, thereby protecting cells from malignant neoplasms (Biegging et al., 2014). The minor polymorphic allele G rs1042522 reduces apoptotic activity, increases the risk of leukemia, and poor overall survival in patients with ALL (Biegging et al., 2014; Tian et al., 2016). The main reason for the contradictory results may be the small sample size and the different ethnicity of the subjects, which indicates the need for replicative genotypic studies of ALL in each specific ethnic population. The lowest population frequency of the minor G allele in Kazakhs – 28.3% compared to previously studied European and Asian populations suggests a relatively low-genetic contribution of TP53 rs1042522 to the risk of developing childhood B-cell ALL, chemotherapy resistance, and the high rates of relapse in the Kazakh population.

As can be seen from Table 4, the frequency of the minor allele G CBR3 rs1056892, which encodes a phase I enzyme of xenobiotic biotransformation, in the Kazakh population was 38.6%, had no significant differences with similar indicators of Europeans – 35.4% and East Asians – 40.5% ($p > 0.05$). Significant differences were found in a higher population frequency of the minor A allele in the South Asians – 53.2% ($p < 0.001$). It has been shown that CBR3 rs1056892 is significantly associated with a 3-fold risk of B-cell ALL with homozygous genotype G/G (OR = 2.77, 95% CI: 1.384-5.565 and $p = 0.004$), the normal homozygous genotype A/A has considerable protective association with B-cell ALL (OR = 0.52, 95% CI: 0.264-1.030 and $p = 0.05$) (Gándara-Mireles et al., 2021). Most of the conducted GWAS have found that the G/G rs1056892 correlates with a 3.3-fold increased risk of cardiomyopathy after anthracycline therapy compared with G/A, A/A genotypes (Lang et al., 2021). Prediction of individual risk and early detection of cardiotoxicity caused by chemotherapy is crucial to prevent irreversible cardiac dysfunction in children with B-cell ALL treatment. ($p > 0.05$).

As shown in Table 4, the minor allele T rs1048943 at *CYP1A1* gene frequency in the studied Kazakh population was 18.2%, which did not differ from its frequency in East Asians – 25.2% ($p > 0.05$) but was significantly higher than the same indicator of Europeans – 3.5% and South Asians – 12.7% ($P < 0.001$). The frequency of the minor allele A rs4646903 at *CYP1A1* gene in the Kazakh population – 33.2% was significantly higher than in the European population – 10.7%, but considerably lower in the East Asian population – 43.0% ($P < 0.001$). No significant differences were found for South Asians – 33.9% ($p > 0.05$).

Figure 1 presents a comparative analysis of the population frequencies of minor alleles 11 SNPs, GWAS associated with B-cell ALL, in the Kazakh population with the studied populations of the world. According to the studied SNPs immune regulation: *GATA3* (rs3824662); transcription and differentiation of B-cells: *ARID5B* (rs7089424, rs10740055), *IKZF1* (rs4132601); differentiation of hematopoietic cells: *PIP4K2A* (rs7088318); apoptosis: *CEBPE* (rs2239633),

tumor suppressors: *CDKN2A* (rs3731249), TP53 (rs1042522); carcinogen metabolism: *CBR3* (rs1056892), *CYP1A1* (rs104894, rs4646903) Kazakhs occupies an intermediate position in comparison with populations of Europe, East and South Asia described in Ensembl genome database project (1,000 Genomes Project, 2021).

Discussion

The results of the comparative analysis indicate that the frequencies of minor alleles of the studied SNPs *ARID5B* (rs7089424, rs10740055), *PIP4K2A* (rs7088318), *CBR3* (rs1056892) in the Kazakh population are closest to Europeans and East Asians, but significantly differ from South Asians. Significant differences are observed for the minor alleles frequency of *GATA3* (rs3824662), *IKZF1* (rs4132601), and *CDKN2A* (rs3731249) ($p < 0.05$).

The minor allele *CEBPE* rs2239633 frequency in the Kazakh population significantly differs only from the Europeans, according to the *CYP1A1* rs104894 from Europeans and South Asians, *CYP1A1* rs4646903 from Europeans and East Asians, according to *TP53* rs1042522 from East and South Asia populations. The results indicate a high genetic heterogeneity of the studied SNPs, which reflects the specific features of Kazakhs formed due to complex evolutionary and migration processes, including the median geographical position between Asia and Europe.

The non-European population is practically not represented in GWAS (Bustamante et al., 2011; Rosenberg et al., 2010). If B-cell ALL is significantly common in non-European populations, GWAS in these ethnic groups has a greater sensitivity to detect new loci. Future GWAS with larger samples of non-Europeans is necessary for a comprehensive characterization of the genetic polymorphisms that predispose to this most common childhood cancer and for the understanding basis of ethnic differences in ALL.

GWAS for childhood B-cell ALL in different populations identified similarities and differences in genetic predisposition to multifactorial disease among ethnic groups. For the first time on a representative population sample of 1,800 Kazakhs, the frequencies distribution of genes according to GWAS associated with the risk of developing childhood B-ALL were studied: immune regulation: *GATA3* (rs3824662); transcription and differentiation of B-cells: *ARID5B* (rs7089424, rs10740055), *IKZF1* (rs4132601); differentiation of hematopoietic cells: *PIP4K2A* (rs7088318); apoptosis: *CEBPE* (rs2239633), tumor suppressors: *CDKN2A* (rs3731249), TP53 (rs1042522); carcinogen metabolism: *CBR3* (rs1056892), *CYP1A1* (rs104894, rs4646903). We conducted a comparative analysis of a genome-wide study the results revealed an intermediate carrier frequency of minor alleles in the Kazakh population between European and Asian for *IKZF1* rs4132601, *PIP4K2A* rs7088318, *CDKN2A* rs3731249, and *CYP1A1* rs4646903. Compared with previously studied populations, Kazakhs have a significantly higher carrying frequency of minor alleles in *GATA3* rs3824662, *ARID5B* rs10740055, and lower carrying frequency in *ARID5B* rs708942,

CEBPE rs2239633, *CYP1A1* rs1048943. MAFs at *TP53* rs1042522 and *CBR3* rs1056892 in the Kazakh population did not differ from the similar indicator in Europeans ($p>0.05$) but were significantly lower than their frequency in Asians ($p<0.001$). So far as the influence of SNPs on the risk of B-cell ALL depends on ethnic factors, differences in LD of genes, lifestyle, and environmental exposure, the results of association studies obtained in one population should be confirmed by replicative genotyping in other ethnic groups.

Therefore, the analysis of the population features of the SNPs frequencies distribution, according to the GWAS associated with the risk of childhood B-cell ALL in the Kazakh population showed:

1. In the Kazakh population, the genotype distribution of 11 SNPs, GWAS associated with the risk of childhood B-cell ALL, is under the Hardy-Weinberg equilibrium ($p>0.05$) obtained as a result of statistical processing in the PLINK-HWE test program.

2. The results of the multiethnic GWAS studies of the search for associations with B-cell ALL, the high population frequencies of minor alleles found in immune regulation genes – *GATA3* rs3824662 – 42.5%; genes of transcription and differentiation of B-cells – *ARID5B* rs7089424 and rs10740055 – 33.1% and 48.5% respectively, suggest their leading significant genetic contribution to the risk of development and prognosis of the effectiveness of B-cell ALL therapy in the Kazakh population.

3. Significantly lower population frequency of the minor allele G *CBR3* rs1056892 – 38.6%, high frequency of the “wild type” allele A – 61.4%, homozygous A/A genotype – 38.6% in Kazakhs suggest a significant protective genetic contribution to the risk of childhood B-cell ALL and fewer cardiac complications after anthracycline therapy.

4. The conducted genome-wide analysis of the Kazakh population showed a significant contribution of hematopoietic cell differentiation genes (rs7088318 polymorphism of the *PIP4K2A* gene); apoptosis (*rs2239633* polymorphism of the *CEBPE* gene), oncosuppressor (*rs1042522* polymorphism of the *TP53* gene); carcinogen metabolism genes (*rs1048943* and *rs4646903* polymorphisms of the *CYP1A1* gene) to the risk of development and efficacy therapy of children’s B-cell ALL, but proven ethnic stratification indicates the need for further replicative genotyping of these polymorphisms in patients with B-cell ALL of Kazakh nationality.

5. The inconsistency of the results of the available GWAS studies, pronounced ethnic stratification, led to choice of 11 “Asian” and “pan-ethnic” polymorphisms for further replicative genotyping of patients with pediatric B-cell ALL in an independent Kazakh population. The obtained results will serve as a basis for the development of effective methods for predicting the risk of development, early diagnosis and effectiveness of treatment of B-cell ALL in children in the Kazakh population.

Author Contribution Statement

All authors contributed equally in this study.

Acknowledgements

The study was carried out within the framework of the project “Genetic Studies of Preeclampsia in Populations of Central Asia and Europe” (InterPregGen) within the 7th Framework Programme of the European Commission under Grant Agreement No. 282540 and the technological programme “National Programme for the Introduction of Personalised and Preventive Medicine in the Republic of Kazakhstan”, OR12165486, ID NCT05121415.

Conflicts of interest

The authors have declared no conflicts of interest for this review.

References

- 1,000 Genomes Project (2021). www.1000genomes.org.
- Agha A, Shabaan H, Abdel-Gawad E, et al (2013). Polymorphism of *CYP1A1* gene and susceptibility to childhood acute lymphoblastic leukemia in Egypt. *Leuk Lymphoma*, **55**, 618-23.
- Akasaka T, Balasas T, Russell LJ, et al (2007). Five members of the CEBP transcription factor family are targeted by recurrent IGH translocations in B-cell precursor acute lymphoblastic leukemia (BCP-ALL). *Blood*, **109**, 3451-61.
- Aydin-Sayitoglu M, Hatirnaz O, Erensoy N, et al (2006). Role of *CYP2D6*, *CYP1A1*, *CYP2E1*, *GSTT1*, and *GSTM1* genes in the susceptibility to acute leukemias. *Am J Hematol*, **81**, 162-70.
- Belson M, Kingsley B, Holmes A (2007). Risk factors for acute leukemia in children: A review. *Environ Health Perspect*, **115**, 138-45.
- Bennett JM, Catovsky D, Daniel M-T, et al (1976). Proposals for the classification of the acute leukaemias. French-American-British (FAB) Co-Operative Group. *Br J Haematol*, **33**, 451-8.
- Berka K, Hendrychová T, Anzenbacher P, et al (2011). Membrane position of ibuprofen agrees with suggested access path entrance to cytochrome P450 2C9 active site. *J Phys Chem*, **115**, 11248-55.
- Berndt SI (2013). Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet*, **45**, 501-12.
- Bieging KT, Mello SS, Attardi LD (2014). Unravelling mechanisms of p53-mediated tumour suppression. *Nat Rev Cancer*, **14**, 359-730.
- Brisson GD, Alves LR, Pombo-de-Oliveira MS (2015). Genetic susceptibility in childhood acute leukaemias: a systematic review. *Cancer Med Sci*, **9**, 539.
- Bustamante CD, Burchard EG, de la Vega FM (2011). Genomics for the world. *Nat*, **475**, 163-5.
- Canalle R, Burim RV, Tone LG, et al (2004). Genetic polymorphisms and susceptibility to childhood acute lymphoblastic leukemia. *Environ Mol Mutagen*, **43**, 100-9.
- CDKN2A* cyclin dependent kinase inhibitor 2A [Homo sapiens (human)] – Gene – NCBI (2021). <https://www.ncbi.nlm.nih.gov/gene/1029>.
- Chen HC, Hu WX, Liu QX, et al (2008). Genetic polymorphisms of metabolic enzymes *CYP1A1*, *CYP2D6*, *GSTM1* and *GSTT1* and leukemia susceptibility. *Eur J Cancer Prev*, **17**, 251-8.
- Clinical Trials Gov (2021). <http://clinicaltrials.gov/>.
- de Smith AJ, Walsh KM, Morimoto LM, et al (2019). Heritable variation at the chromosome 21 gene *ERG* is associated with acute lymphoblastic leukemia risk in children with and
- Asian Pacific Journal of Cancer Prevention, Vol 24 2663*

- without Down syndrome. *Leuk*, **33**, 2746-51.
- Do TN, Ucisik-Akkaya E, Davis CF, et al (2009). TP53 R72P and MDM2 SNP309 polymorphisms in modification of childhood acute lymphoblastic leukemia susceptibility. *Cancer Genet Cytogenet*, **195**, 31-6.
- Ellinghaus E, Stanulla M, Richter G, et al (2012). Identification of germline susceptibility loci in ETV6-RUNX1-rearranged childhood acute lymphoblastic leukemia. *Leuk*, **26**, 902-9.
- Ensembl (2021). <http://asia.ensembl.org/index.html>.
- ExAC Browser (2021). <https://gnomad.broadinstitute.org/>.
- Fiume R, Stijf-Bultsma Y, Shah ZH, et al (2015). PIP4K and the role of nuclear phosphoinositides in tumour suppression. *BBA – Mol Cell Biol Lipids*, **1851**, 898-910.
- Gándara-Mireles JA, Lares-Asseff I, Espinoza EAR, et al (2021). Genotype Analysis of ABCC1, NCF4 and CBR3 Polymorphism and the Association with Childhood Acute Lymphoblastic Leukemia in Mexican Childhood Population. *Front Pharmacol*, **11**, 616630.
- Gharbi H, Hassine IB, Soltani I, et al (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**, 157-67.
- GnomAD (2021). <https://gnomad.broadinstitute.org>.
- Górniak P, Pastorczak A, Zalewska-Szewczyk B, et al (2014). Polymorphism in *IKZF1* gene affects age at onset of childhood acute lymphoblastic leukemia. *Leuk Lymphoma*, **55**, 2174-8.
- Greaves M (2006). Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer*, **6**, 193-203.
- Gutierrez-Camino A, Martin-Guerrero I, de Andoin NG, et al (2017). Confirmation of involvement of new variants at *CDKN2A/B* in pediatric acute lymphoblastic leukemia susceptibility in the Spanish population. *PLoS One*, **12**, e0177421.
- GWAS Catalog (2021). https://www.ebi.ac.uk/gwas/efotraits/EFO_0000094.
- HapMap (2021). <http://hapmap.ncbi.nlm.nih.gov/index.html.en>.
- Harris JM, Jaffe ES, Dieblod J, et al (1999). World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting-Airlie House, Virginia, November 1997. *J Clin Oncol*, **17**, 3835-49.
- Hirschhorn JN, Gajdos ZK (2011). Genome-wide association studies: results from the first few years and potential implications for clinical medicine. *Ann Rev Med*, **62**, 11-24.
- Homo sapiens GATA binding protein 3 (*GATA3*), transcript variant 2, mRNA – Nucleotide – NCBI 2021. https://www.ncbi.nlm.nih.gov/nuccore/NM_001002295.1.
- Homo sapiens GATA binding protein 3 2021. <https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=2625>.
- Hosgood HD, Au W-Y, Kim HN, et al (2013). IL10 and TNF variants and risk of non-Hodgkin lymphoma among three Asian populations. *Int J Hematol*, **97**, 793-9.
- Joseph T, Kusumakumary P, Chacko P, et al (2004). Genetic polymorphism of CYP1A1, CYP2D6, GSTM1 and GSTT1 and susceptibility to acute lymphoblastic leukaemia in Indian children. *Pediatr Blood Cancer*, **43**, 560-7.
- Kadan-Lottick NS, Ness KK, Bhatia S, Gurney JG (2003). Survival variability by race and ethnicity in childhood acute lymphoblastic leukemia. *JAMA*, **290**, 2008-14.
- Kampouraki E, Lourou M, Zervou MI, et al (2021). Zervou et al Role of CXCL12, TP53 and CYP1A1 gene polymorphisms in susceptibility to pediatric acute lymphoblastic leukemia. *Oncol Lett*, **22**, 659.
- Kano G, Morimoto A, Takashi M, et al (2008). Ikaros dominant negative isoform (Ik6) induces IL-3-independent survival of murine pro-B lymphocytes by activating JAK-STAT and up-regulating Bcl-x1 levels. *Leuk Lymphoma*, **49**, 965-73.
- Krajinovic M, Labuda D, Mathonnet G, et al (2002). Polymorphisms in genes encoding drugs and xenobiotic metabolizing enzymes, DNA repair enzymes, and response to treatment of childhood acute lymphoblastic leukemia. *Clin Cancer Res*, **8**, 802-10.
- Krajinovic M, Labuda D, Richer C, et al (1999). Susceptibility to childhood acute lymphoblastic leukemia: influence of CYP1A1, CYP2D6, GSTM1, and GSTT1 genetic polymorphisms. *Blood*, **93**, 1496-1501.
- Kreile M, Piekuse L, Rots D, et al (2016). Analysis of possible genetic risk factors contributing to development of childhood acute lymphoblastic leukaemia in the Latvian population. *Arch Med Sci*, **12**, 479-85.
- Lakhman SS, Ghosh D, Blanco JG (2005). Functional significance of a natural allelic variant of human carbonyl reductase 3 (CBR3). *Drug Metab Dispos*, **33**, 254-7.
- Lang JK, Karthikeyan B, Quiñones-Lombraña A, et al (2021). CBR3 V244M is associated with LVEF reduction in breast cancer patients treated with doxorubicin. *Cardio-Oncol*, **7**, 17.
- Lautner-Csorba O, Gézsi A, Semsei AF (2012). Candidate gene association study in pediatric acute lymphoblastic leukemia evaluated by Bayesian network based Bayesian multilevel analysis of relevance. *BMC Med Genom*, **5**, 42.
- Li S, Ren L, Fan L, et al (2015). *IKZF1* rs4132601 polymorphism and acute lymphoblastic leukemia susceptibility: a meta-analysis. *Leuk Lymphoma*, **56**, 978-82.
- Lin CY, Li MJ, Chang JG, et al (2013). 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**, 140-5.
- Linabery AM, Ross JA (2008). Trends in childhood cancer incidence in the U.S. (1992-2004). *Cancer*, **112**, 416-32.
- Linkage Disequilibrium Calculator (2021). <https://asia.ensembl.org/Help/View?id=587>.
- Liu J, Weiling G, Xueqin L, et al (2021). The *CEBPE* rs2239633 genetic polymorphism on susceptibility to childhood acute lymphoblastic leukemia: an updated meta-analysis. *Environ Health Prev Med*, **26**, 2.
- Lopez-Lopez E, Gutierrez-Camino A, Martin-Guerrero I, et al (2013). Re: novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. *JNCI*, **105**, 1512-3.
- Lu J, Zhao Q, Zhai Y, et al (2015). Genetic polymorphisms of CYP1A1 and risk of leukemia: a meta-analysis. *Oncotargets Ther*, **8**, 2883-2902.
- Mahjoub S, Ghali R, Chaieb V, et al (2018). *CDKN2A* and *CDKN2B* Gene Variants in Acute Lymphoblastic Leukemia in Tunisian Population. *J Leuk*, **6**, 1000247.
- Manning AK (2012). A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet*, **44**, 659-69.
- Migliorini G, Fiege B, Hosking FJ, et al (2013). Variation at 10p12.2 and 10p14 influences risk of childhood B-cell acute lymphoblastic leukemia and phenotype. *Blood*, **122**, 3298-3307.
- Moriyama T, Metzger ML, Wu G, et al (2015). Germline genetic variation in ETV6 and risk of childhood acute lymphoblastic leukaemia: a systematic genetic study. *Lancet Oncol*, **16**, 1659-66.
- Mosad Y, Baraka T, Darwis A, et al (2016). *GATA3* rs3824662 gene polymorphism as possible risk factor in a cohort

- of Egyptian patients with pediatric acute lymphoblastic leukemia and its prognostic impact. *Leuk Lymphoma*, **58**, 689-98.
- Mullighan ChG, Su X, Zhang J, et al (2009). Deletion of *IKZF1* and prognosis in acute lymphoblastic leukemia. *New Eng J Med*, **360**, 470-80.
- Nakajima H, Watanabe N, Shibata F, et al (2006). N-terminal region of CCAAT/enhancer-binding protein epsilon is critical for cell cycle arrest, apoptosis, and functional maturation during myeloid differentiation. *J Biol Chem*, **281**, 14494-502.
- National Center for Biotechnology Information (2021). <http://www.ncbi.nlm.nih.gov>.
- National Human Genome Research Institute (2021). www.genome.gov.
- Nazki FH, Masood A, Banday MA, et al (2012). Thymidylate synthase enhancer region polymorphism not related to susceptibility to acute lymphoblastic leukemia in the Kashmir population. *Genet Mol Res*, **11**, 906-17.
- Nebert DW, McKinnon RA, Puga A (1996). Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer. *DNA Cell Biol*, **15**, 273-80.
- Orsi L, Rudant, J, Bonaventure A, et al (2012). Genetic polymorphisms and childhood acute lymphoblastic leukemia: GWAS of the ESCALE study (SFCE). *Leuk*, **26**, 2561-4.
- Papaemmanuil E, Hosking FJ, Vijaykrishnan J, et al (2009). Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet*, **41**, 1006-10.
- Perez-Andreu V, Roberts KG, Harvey RC, et al (2013). Inherited *GATA3* variants are associated with Ph-like childhood acute lymphoblastic leukemia and risk of relapse. *Nat Genet*, **45**, 1494-8.
- Perez-Andreu V, Roberts KG, Xu H, et al (2015). A genome-wide association study of susceptibility to acute lymphoblastic leukemia in adolescents and young adults. *Blood*, **125**, 680-6.
- Plink (2021). <http://pngu.mgh.harvard.edu/purcell/plink>.
- Prasad RB, Hosking FJ, Vijaykrishnan J, et al (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**, 1765-7.
- Purcell S (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, **81**, 559.
- Qian M, Cao X, Devidas M, et al (2017). TP53 Germline Variations Influence the Predisposition and Prognosis of B-Cell Acute Lymphoblastic Leukemia in Children. *J Clin Oncol*, **36**, 591-9.
- Qian M, Zhao X, Devidas M, et al (2019). Genome-wide association study of susceptibility loci for T-cell acute lymphoblastic leukemia in children. *J Natl Cancer Inst*, **111**, 1350-7.
- Razmkhah F, Pazhakh V, Zaker F, et al (2011). Frequency of CYP1A1*2C polymorphism in patients with leukemia in the Iranian population. *Lab Med*, **42**, 220-3.
- Rosenberg NA, Huang L, Jewett EM (2010). Genome-wide association studies in diverse populations. *Nat Rev Genet*, **11**, 356-66.
- Rudant J, Orsi L, Bonaventure A, et al (2015). *ARID5B*, *IKZF1* and non-genetic factors in the etiology of childhood acute lymphoblastic leukemia: The ESCALE study. *PLoS One*, **10**, e0121348.
- Sellers M, Kastner P, Chan S (2011). Ikaros in B cell development and function. *World J Biol Chem*, **2**, 132-9.
- Sherborne AL, Hosking FJ, Prasad RB, et al (2010). Variation in *CDKN2A* at 9p21.3 influences childhood acute lymphoblastic leukemia risk. *Nat Genet*, **42**, 492-4.
- Shin YJ, Sa JK, Lee Y, et al (2019). *PIP4K2A* as a Negative Regulator of PI3K in PTEN-Deficient Glioblastoma. *J Exp Med*, **216**, 1120-34.
- Song C, Gowda C, Pan X (2015). Targeting casein kinase II restores Ikaros tumor suppressor activity and demonstrates therapeutic efficacy in high-risk leukemia. *Blood*, **126**, 1813-22.
- Speedy HE, Di Bernardo MC, Sava GP, et al (2014). A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. *Nat Genet*, **46**, 56-60.
- Sun J, Zheng, J, Tang L, et al (2015). Association between *CEBPE* variant and childhood acute leukemia risk: evidence from a meta-analysis of 22 studies. *PLOS One*, **10**, e0125657.
- Swinney RM, Beuten J, Collier AB, et al (2011). Polymorphisms in *CYP1A1* and ethnic-specific susceptibility to acute lymphoblastic leukemia in children. *Cancer Epidemiol, Biomarkers Prev*, **20**, 1537-42.
- The Database of Genotypes and Phenotypes (2021). <https://www.ncbi.nlm.nih.gov/gap/>.
- The integrated Cancer Repository for Cancer Research (iCaRe2) (2021). <https://www.unmc.edu/cancercenter/registry/>.
- Tian X, Dai S, Sun J, et al (2016). Association between TP53 Arg72Pro polymorphism and leukemia risk: A meta-analysis of 14 case-control studies. *Sci Rep*, **6**, 24097.
- Trevino LR, Yang W, French D, et al (2009). Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nat Genet*, **41**, 1001-5.
- Vijaykrishnan J, Qian M, Yang JJ (2019). Identification of four novel associations for B-cell acute lymphoblastic leukaemia risk. *Nat Commun*, **10**, 5348.
- Vijaykrishnan J, Sherborne AL, Sawangpanich R, et al (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**, 1870-4.
- Vijaykrishnan J, Studd J, Broderick P, et al (2018). Genome-wide association study identifies susceptibility loci for B-cell childhood acute lymphoblastic leukemia. *Nat Commun*, **9**, 1340.
- Walsh KM, Chokkalingam AP, Hsu LI, et al (2013). Associations between genome-wide Native American ancestry, known risk alleles and B-cell ALL risk in Hispanic children. *Leuk*, **27**, 2416-19.
- Walsh KM, de Smith AJ, Hansen HM, et al (2015). A heritable missense polymorphism in *CDKN2A* confers strong risk of childhood acute lymphoblastic leukemia and is preferentially selected during clonal evolution. *Cancer Res*, **75**, 4884-94.
- Wang Y, Chen J, Li J, et al (2013). Association of three polymorphisms in *ARID5B*, *IKZF1* and *CEBPE* with the risk of childhood acute lymphoblastic leukemia in a Chinese population. *Gene*, **524**, 203-7.
- Wiemels JL, Walsh KM, de Smith AJ, et al (2018). GWAS in childhood acute lymphoblastic leukemia reveals novel genetic associations at chromosomes 17q12 and 8q24.21. *Nat Commun*, **9**, 286.
- Wilsker D, Patsialou A, Dallas PB, et al (2002). ARID proteins: a diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development. *Nat Library Med*, **13**, 95-106.
- Wu XC, Andrews PA, Chen VW, et al (2009). Incidence of extranodal non-Hodgkin lymphomas among whites, blacks, and Asians/Pacific Islanders in the United States: anatomic site and histology difference. *Cancer Epidemiol*, **2009**, 337-46.
- Xu H, Yang W, Perez-Andreu V, et al (2013). Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic

- leukemia in ethnically diverse populations. *J Nat Cancer Inst*, **105**, 733-42.
- Xu H, Zhang H, Yang W, et al (2015). Inherited coding variants at the *CDKN2A* locus influence susceptibility to acute lymphoblastic leukaemia in children. *Nat Commun*, **6**, 7553.
- Xu K, Feng Q, Wiemels JL, et al (2021). Disparities in acute lymphoblastic leukemia risk and survival across the lifespan in the United States of America. *J Transl Genet Genomics*, **5**, 218-39.
- Yamamoto JF, Goodman MT (2008). Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002. *Cancer Causes Control*, **19**, 379-90.
- Yang JJ, Cheng C, Devidas M, et al (2011). Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. *Nat Genet*, **43**, 237-41.
- Zhang H, Pak-Yin Liu A, Meenakshi D, et al (2020). Association of *GATA3* polymorphisms with minimal residual disease and relapse risk in childhood acute lymphoblastic leukemia. *J Nat Cancer Inst*, **113**, 408-17.
- Zhang S, Li Z, Yan X, et al (2019). Regulatory Network and Prognostic Effect Investigation of *PIP4K2A* in Leukemia and Solid Cancers. *Front Genet*, **9**, 721.
- Zhuo W, Zhang L, Zhu B, et al (2012). Association between CYP1A1 Ile462Val variation and acute leukemia risk: meta-analyses. *PLoS One*, **7**, e46974.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.