

Article

Identification of Genomic Variants Associated with the Risk of Acute Lymphoblastic Leukemia in Native Americans from Brazilian Amazonia

Luciana P. C. Leitão ^{1,2,†}, Darlen C. de Carvalho ^{1,†}, Juliana C. G. Rodrigues ¹, Marianne R. Fernandes ¹, Alayde V. Wanderley ³, Lui W. M. S. Vinagre ¹, Natasha M. da Silva ¹, Lucas F. Pastana ¹, Laura P. A. Gellen ¹, Matheus C. E. Assunção ¹, Sweny S. M. Fernandes ³, Esdras E. B. Pereira ^{4,5}, André M. Ribeiro-Dos-Santos ⁴, João F. Guerreiro ⁴, Ândrea Ribeiro-dos-Santos ⁴, Paulo P. de Assumpção ¹, Sidney E. B. dos Santos ⁴ and Ney P. C. dos Santos ^{1,*}

- ¹ Oncology Research Nucleus, Universidade Federal do Pará, Belém 66073-005, PA, Brazil; colaresluciana@gmail.com (L.P.C.L.); darlen.c.carvalho@gmail.com (D.C.d.C.); julianacgrodriques@gmail.com (J.C.G.R.); fernandesmr@yahoo.com.br (M.R.F.); luivinagre@gmail.com (L.W.M.S.V.); ntshmonte@gmail.com (N.M.d.S.); lucas.favacho.celular@gmail.com (L.F.P.); laura.patricia.agellen@hotmail.com (L.P.A.G.); matheus.assuncao@icb.ufpa.br (M.C.E.A.); assumppcaopp@gmail.com (P.P.d.A.)
- ² Faculdade de Ciências Médicas do Pará (FACIMPA), Marabá 68508-030, PA, Brazil
- ³ Pediatrics Department, Ophir Loyola Hospital, Belém 66063-240, PA, Brazil; alaydevieira@yahoo.com.br (A.V.W.); swenymf@gmail.com (S.S.M.F.)
- ⁴ Human and Medical Genetics Laboratory, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém 66075-110, PA, Brazil; esdrasedgarbp@gmail.com (E.E.B.P.); andremrsantos@gmail.com (A.M.R.-D.-S.); joao.guerreiro53@gmail.com (J.F.G.); akely@ufpa.br (Â.R.-d.-S.); sidneysantos@ufpa.br (S.E.B.d.S.)
- ⁵ Instituto Tocantinense Presidente Antônio Carlos (ITPAC), Abaetetuba 68440-000, PA, Brazil
- * Correspondence: npcsantos.ufpa@gmail.com; Tel.: +55-(91)-3201-6778
- † These authors contributed equally to this work.



Citation: Leitão, L.P.C.; de Carvalho, D.C.; Rodrigues, J.C.G.; Fernandes, M.R.; Wanderley, A.V.; Vinagre, L.W.M.S.; da Silva, N.M.; Pastana, L.F.; Gellen, L.P.A.; Assunção, M.C.E.; et al. Identification of Genomic Variants Associated with the Risk of Acute Lymphoblastic Leukemia in Native Americans from Brazilian Amazonia. *J. Pers. Med.* **2022**, *12*, 856. <https://doi.org/10.3390/jpm12060856>

Academic Editor: Sergiu Pasca

Received: 18 March 2022

Accepted: 18 April 2022

Published: 25 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: A number of genomic variants related to native American ancestry may be associated with an increased risk of developing Acute Lymphoblastic Leukemia (ALL), which means that Latin American and hispanic populations from the New World may be relatively susceptible to this disease. However, there has not yet been any comprehensive investigation of the variants associated with susceptibility to ALL in traditional Amerindian populations from Brazilian Amazonia. We investigated the exomes of the 18 principal genes associated with susceptibility to ALL in samples of 64 Amerindians from this region, including cancer-free individuals and patients with ALL. We compared the findings with the data on populations representing five continents available in the 1000 Genomes database. The variation in the allele frequencies found between the different groups was evaluated using Fisher's exact test. The analyses of the exomes of the Brazilian Amerindians identified 125 variants, seven of which were new. The comparison of the allele frequencies between the two Amerindian groups analyzed in the present study (ALL patients vs. cancer-free individuals) identified six variants (rs11515, rs2765997, rs1053454, rs8068981, rs3764342, and rs2304465) that may be associated with susceptibility to ALL. These findings contribute to the identification of genetic variants that represent a potential risk for ALL in Amazonian Amerindian populations and might favor precision oncology measures.

Keywords: Acute Lymphoblastic Leukemia; Amerindian populations; genetic susceptibility

1. Introduction

Acute Lymphoblastic Leukemia (ALL) is the most common cancer in children and is the principal cause of childhood mortality due to malignant disease [1,2]. The genetic etiology of ALL is driven by an ample diversity of alterations of the pathways responsible for the regulation of the cell cycle in the lymphoid precursors of the B and T cell lines, which

include chromosomal translocations, mutations, and aneuploidy [3–5]. In recent years, Genome-Wide Association Studies (GWASs) have identified a number of loci associated with the risk of developing ALL, including the *ARID5B*, *IKZF1*, *PIP4K2A*, *CEBPE*, *GATA3*, *BMI*, and *CDKN2A* genes [6–11].

Most GWASs addressing ALL susceptibility have focused on homogeneous populations in regions such as Europe or North America, whereas the highest incidence of childhood ALL is found in populations with a major component of native American ancestry, such as the Latin American and hispanic populations of the New World [9,12–14]. These populations are highly admixed and, like the population of Brazil, are primarily descended from European, African, and native American ancestors [15,16], and the high incidence of ALL observed in Latin American and hispanic populations has been attributed [9,12,13,17] to genetic risk factors related to native American ancestry. Despite the evidence that genetic variants related to native American ancestry may influence the incidence of childhood ALL, no data are available on the distribution of these variants in traditional Amerindian populations.

Given this, in the present study, we investigated the genetic variants potentially involved in the etiology of ALL in traditional Amerindian populations from Brazilian Amazonia. For this, we used New Generation Sequencing (NGS) to define the exomes of 17 genes associated with susceptibility to ALL in samples obtained from indigenous groups that inhabit Brazilian Amazonia, in both ALL patients and cancer-free individuals. The variants encountered in this initial analysis were compared with the data available on populations representing five different continents, which were obtained from the 1000 Genomes database.

2. Materials and Methods

2.1. Ethics, Consent, and Permissions

The present study was approved by Brazilian National Committee on Research Ethics—CONEP (identified by No 1062/2006 and 123/1998). All participants signed a free-informed consent as well as the tribe leaders and when necessary, a translator explained the project and the importance of the research. Their materials were collected according to the Declaration of Helsinki.

2.2. Study Population

The study sample included 59 cancer-free Amerindians and five individuals diagnosed with ALL. All the participants were members of isolated ethnic groups located in Brazilian Amazonia (Table S1). The genomic Amerindian ancestry of all these individuals was quantified and found to be at least 64% in all cases. The Amerindians with ALL were diagnosed and treated in two public hospitals specialized in the treatment of childhood cancer, the Ophir Loyola Hospital and the Octavio Lobo Childhood Oncology Hospital, both located in the city of Belém, in Pará state, northern Brazil. The clinical and demographic data on these patients are presented in Table S2.

Data were obtained from the 1000 Genomes Project database (available at <https://www.1000genomes.org>, accessed on: 20 February 2021) to provide comparisons with ethnic groups from other continents. This sample included 661 individuals from Africa (AFR), 503 from Europe (EUR), 347 from the Americas (AMR), 504 from East Asia (EAS), and 489 from South Asia (SAS).

2.3. Selection of the Genes

A total of 17 genes were selected for the present study (see Table S3). The genes were selected based on a search of the literature available in the NCBI and Ensembl databases, and on GWASs or studies of other genetic markers associated with the risk of ALL. More details are described in Table S4.

2.4. Extraction of the DNA and Preparation of the Exomes

Samples of 5 mL of peripheral blood were collected from each of the participants of the study. The genetic material was extracted from these blood samples using the Roche Applied Science DNA extraction kit (Roche, Penzberg, Germany) following the manufacturer's instructions, and it was quantified using a NanoDrop 1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). The exome libraries were prepared using the commercial Nextera Rapid Capture Exome kit (Illumina, San Diego, CA, EUA) and the SureSelect Human All Exon V6 kit (Agilent, Santa Clara, CA, USA), with the manufacturer's protocol being followed in both cases. The sequencing reactions were run in the NextSeq 500[®] platform (Illumina[®], San Diego, CA, USA) using the NextSeq 500 high-output v2 300 cycle kit (Illumina[®], San Diego, CA, USA).

2.5. Bioinformatic Analysis

The bioinformatic analyses followed the approach described by Ribeiro-Dos-Santos et al. [18] and Rodrigues et al. [19]. For this, the sequences were first filtered to eliminate low-quality reads, and then mapped and aligned with the reference genome (GRCh38) using BWA v.0.7. The alignment was then processed to remove duplicate sequences, recalibrate the mapping quality, and finalize local realignment. The results were processed in GATK v.3.2 to identify the reference genome variants. The Viewer of Variants (ViVa[®], Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil) software was used to analyze the annotations of the variants. The variants were annotated in three databases—SnEff v.4.3.T, Ensembl Variant Effect Predictor (Ensembl version 99), and ClinVar (v.2018–10). The SIFT (v.6.2.1), PolyPhen-2 (v.2.2), LRT (November, 2009), Mutation Assessor (v.3.0), Mutation Taster (v.2.0), FATHMM (v.2.3), PROVEAN (v.1.1.3), MetaSVM (v.1.0), M-CAP (v.1.4), and FATHMM-MKL databases were used for the in silico prediction of pathogenicity.

2.6. Statistical Analysis

For the statistical analyses, the cancer-free Amerindians were assigned to the Native (NAT) group, while the ALL patients were in the ALL_NAT (Native with ALL) group. All the analyses were run in the R v.3.5.1 program. The differences in the allelic frequencies between the NAT and ALL_NAT were evaluated using Fisher's exact test, which was also applied to the comparisons with the continental populations. A $p \leq 0.05$ significance level was considered for all the analyses.

3. Results

The analysis of the exomes of the 64 Amerindians investigated in the present study revealed the presence of 125 variants, seven of which were new. The SnEff software [20] was used to annotate and predict the effects of these variants. This procedure classifies the impact of the variants in four categories: (i) modifier (no evidence of impact), (ii) low (no apparent alteration of protein function), (iii) moderate (some alteration of protein function), and (iv) high (high level of impact on protein function). The 118 variants identified in the 17 genes investigated in the present study are described in Table S3. Four of the new variants were classified as modifiers (Table 1), two as low effect, and one as moderate. The majority of these new variants were found in the ALL_NAT group, at frequencies of less than 0.1%.

As some of the variants were not covered adequately in one or other of the groups (NAT or ALL_NAT), they were excluded from the comparison of allele frequencies. This left 64 variants that were included in the analysis of association. The frequencies of these variants are compared between groups in Table 2.

Table 1. The new gene variants identified in the Amerindian population investigated in the present study.

Chromosome	Chromosomal Position	Gene	Impact	Reference Allele	Group
chr7	50368329	<i>IKZF1</i>	MODIFIER	C	ALL_NAT
chr16	78108386	<i>WWOX</i>	MODIFIER	T	ALL_NAT
chr16	78386853	<i>WWOX</i>	LOW	A	NAT
chr17	39832190	<i>IKZF3</i>	MODIFIER	T	NAT
chr17	39765793	<i>IKZF3</i>	MODERATE	C	ALL_NAT
chr17	39868422	<i>ZPBP2</i>	MODIFIER	C	ALL_NAT
chr17	39765799	<i>IKZF3</i>	LOW	C	ALL_NAT

Table 2. Comparisons of the frequencies of the alleles investigated in the present study between the two Amerindian groups, that is, ALL patients (ALL_NAT) and cancer-free individuals (NAT).

dbSNP	Gene	Frequency in Group:			* <i>p</i>
		ALL_NAT	NAT		
rs773061413	<i>ARID5B</i>	0.20	0	0.0781	
rs513349	<i>BAK1</i>	0.80	0.56	0.3868	
rs11515	<i>CDKN2A</i>	0.60	1	0.0050	
rs3088440	<i>CDKN2A</i>	0.50	0.56	1	
rs974336	<i>CDKN2B</i>	0.25	0	0.0781	
rs2302901	<i>ELK3</i>	0	0.03	1	
rs422628	<i>GATA3</i>	0.75	0.97	0.2197	
rs2305479	<i>GSDMB</i>	0.30	0.43	1	
rs2305480	<i>GSDMB</i>	0.30	0.43	1	
rs11078928	<i>GSDMB</i>	0.30	0.43	1	
rs12450091	<i>GSDMB</i>	0.20	0.26	1	
rs11078927	<i>GSDMB</i>	0.30	0.42	1	
rs8068981	<i>IGF2BP1</i>	0.60	0	0.0002	
rs62078405	<i>IGF2BP1</i>	0.20	0	0.0781	
rs2289637	<i>IGF2BP1</i>	0.20	0.06	0.3434	
rs10899750	<i>IKZF1</i>	0.50	0.67	0.6572	
rs61731355	<i>IKZF1</i>	0.20	0.31	1	
rs12669559	<i>IKZF1</i>	0.10	0.28	1	
rs907092	<i>IKZF3</i>	0.20	0.04	0.2197	
rs3824810	<i>LHPP</i>	0.60	0.57	1	
rs3824809	<i>LHPP</i>	0.60	0.71	0.6287	
rs6597801	<i>LHPP</i>	0.70	0.88	0.1907	
rs943192	<i>PIP4K2A</i>	0.80	0.98	0.1513	
rs2765997	<i>PIP4K2A</i>	0.50	0.96	0.0043	
rs1132816	<i>PIP4K2A</i>	0.10	0.01	0.1774	
rs61731109	<i>PIP4K2A</i>	0.10	0.01	0.1774	
rs1053454	<i>PIP4K2A</i>	0.90	0	7.264 × 10⁻⁷	
rs2447919	<i>USP7</i>	0.13	0.31	1	
rs11551182	<i>USP7</i>	0.30	0.40	1	
rs139138924	<i>USP7</i>	1	1	1	
rs2304465	<i>USP7</i>	0.40	0	0.0050	
rs1382390	<i>USP7</i>	1.00	1	1	
rs2447918	<i>USP7</i>	0.16	0	0.0781	
rs117832776	<i>WWOX</i>	0.20	0	0.0781	
rs383362	<i>WWOX</i>	0.30	0.16	0.3052	
rs77897021	<i>WWOX</i>	0.10	0.08	0.4545	
rs12934985	<i>WWOX</i>	0.10	0.21	1	
rs3764342	<i>WWOX</i>	0.40	0	0.0050	
rs200461412	<i>WWOX</i>	0.20	0.03	0.2197	
rs140060332	<i>WWOX</i>	0.20	0	0.0781	
rs4130513	<i>WWOX</i>	0.30	0.09	0.1551	
rs11545029	<i>WWOX</i>	0.10	0.08	0.4545	

Table 2. Cont.

dbSNP	Gene	Frequency in Group:			* <i>p</i>
		ALL_NAT	NAT		
rs8050128	WWOX	0.75	0.41	0.1590	
rs75559202	WWOX	0.10	0.09	0.4545	
rs146481440	WWOX	0.20	0.06	0.3434	
rs12446823	WWOX	0.10	0.01	0.1774	
rs2288034	WWOX	0.60	0.32	0.3292	
rs144601717	WWOX	0.10	0.01	0.1774	
rs2303190	WWOX	0	0.22	0.5739	
rs3764340	WWOX	0.20	0.09	0.3990	
rs2303191	WWOX	0.80	0.98	0.1513	
rs2288035	WWOX	0.10	0.08	0.4545	
rs8048830	WWOX	0.25	0.01	0.1513	
rs67493355	WWOX	0.25	0.16	1	
rs202093359	WWOX	0.20	0.08	0.3990	
rs11545028	WWOX	0.10	0.22	1	
rs2288033	WWOX	0.40	0.25	0.5917	
rs7199110	WWOX	0.60	0.40	0.6423	
rs555396422	WWOX	0.10	0	0.0923	
rs384216	WWOX	0	0	1	
rs75027016	ZPBP2	0.30	0.15	0.2662	
rs11557467	ZPBP2	0.30	0.43	1	
rs11557466	ZPBP2	0.30	0.42	1	
rs10852935	ZPBP2	0.30	0.43	1	

* Fisher’s exact test.

The analyses revealed significant differences between the two groups in six variants of five genes: *PIP4K2A* (variants rs2765997 and rs1053454), *CDKN2A* (rs11515), *IGF2BP1* (rs8068981), *USP7* (rs2304465), and *WWOX* (rs3764342). The allele frequencies of these six variants recorded in the ALL_NAT group were also compared with those recorded in the 1000 Genomes Project for the five continental populations (AFR, AMR, EAS, EUR, and SAS). The comparisons are shown in Table 3 and the *p* values, in Table 4.

The analyses revealed significant differences between the Amerindian group (ALL_NAT) and all the continental populations (AFR, AMR, EAS, EUR and SAS) in the frequency of the rs11515 variant of the *CDKN2A* gene. The frequency of the rs8068981 variant of the *IGF2BP1* gene was also significantly different from that of the EAS population, while the frequency of the rs2765997 variant of the *PIP4K2A* gene was significantly different from those of the AFR and EAS populations.

Table 3. Allele frequencies of the six key variants identified in the present study in the ALL_NAT populations (Table 2) and the five continental populations from the 1000 Genomes Project (AFR, AMR, EAS, EUR, and SAS).

dbSNP	Gene	Frequency in the Population:					SAS
		ALL_NAT	AFR	AMR	EAS	EUR	
rs11515	<i>CDKN2A</i>	0.60	0.16	0.08	0.02	0.15	0.05
rs8068981	<i>IGF2BP1</i>	0.60	0.57	0.68	0.95	0.71	0.82
rs2765997	<i>PIP4K2A</i>	0.50	0.18	0.20	0.05	0.37	0.37
rs1053454	<i>PIP4K2A</i>	0.90	0.85	0.81	0.65	0.64	0.78
rs2304465	<i>USP7</i>	0.40	0.66	0.58	0.70	0.59	0.49
rs3764342	<i>WWOX</i>	0.40	0.55	0.14	0.30	0.11	0.32

Table 4. The p values recorded for the pairwise comparisons (ALL_NAT vs. AFR, AMR, EAS, EUR or SAS) of the allele frequencies of the six key variants identified in the present study (see Table 3).

dbSNP	Gene	Frequency ALL_NAT	p for the ALL_NAT Population Versus:				
			AFR	AMR	EAS	EUR	SAS
rs11515	<i>CDKN2A</i>	0.67	0.0349	0.0050	0.0001	0.0267	0.0015
rs8068981	<i>IGF2BP1</i>	0.60	1	0.6577	0.0263	0.6283	0.2302
rs2765997	<i>PIP4K2A</i>	0.50	0.0452	0.0579	0.0015	0.3627	0.3663
rs1053454	<i>PIP4K2A</i>	0.90	1	0.5880	0.1702	0.1655	0.5911
rs2304465	<i>USP7</i>	0.40	0.3447	0.6535	0.1602	0.6536	1
rs3764342	<i>WWOX</i>	0.40	0.6626	0.1595	0.6411	0.0960	0.6550

4. Discussion

The incidence of ALL is relatively high in populations with a high degree of native American ancestry, such as Latin American and Hispanic populations, which has been attributed, in part, to the contribution of molecular markers associated with a high risk of ALL in these populations [9,12–14,21]. Despite this known association, no genomic data are available on the susceptibility to ALL of traditional Amerindian populations from Brazilian Amazonia. In the present study, we investigated the exomes of 17 of the principal genes associated with susceptibility to ALL in samples of indigenous Amazonian populations, including ALL patients and cancer-free individuals. It would be interesting to demonstrate whether the results observed for the markers investigated here in Amerindian populations would demonstrate the same profile in a larger cohort of leukemic patients from the same ethnic group. However, samples from indigenous patients with ALL are difficult to obtain, given the rarity of the disease and due to the fact that these populations inhabit remote and difficult-to-access rural regions, which reflects on the difficulty of care and clinical follow-up of these patients.

During the study, we identified seven new variants in four genes (*IKZF1*, *IKZF3*, *WWOX*, and *ZPBP2*), most of which were present in the ALL_NAT group, that is, Amerindian ALL patients. These genes are known to have alleles associated with some level of risk in the etiology of ALL in different populations [17,22–24]. Given this, we believe that the new variants identified in the present study should be investigated further as potential risk factors for the incidence of ALL in Amerindian populations.

We also identified six variants of five genes that we associated with susceptibility to ALL (*CDKN2A*_rs11515, *PIP4K2A*_rs2765997 and rs1053454, *IGF2BP1*_rs8068981, *WWOX*_rs3764342, and *USP7*_rs2304465) given the significant differences in the frequencies recorded in the two Amerindian groups (ALL patients and cancer-free individuals), and between the ALL group and populations of other groups from around the world.

The *CDKN2A* gene plays an important role in leukemogenesis. The rs11515 variant is located in the 3'-untranslated region (UTR) of the *CDKN2A* gene, and is known to be associated with a number of different types of cancer [25], including breast cancer [26], glioblastoma [27], melanoma [28], and colorectal cancer [29]. However, no data are available on the possible influence of this variant on the risk of developing ALL. In the present study, we recorded a high frequency of this variant in the Amerindian population, which was significantly higher than that recorded in other populations around the world. This indicates that research on this variant should be prioritized for the identification of its potential role in the etiology of ALL in this population.

A number of previous studies of genetic polymorphisms in the *PIP4K2A* gene have identified an association with susceptibility to ALL [30,31]. We identified an association with ALL in two variants (rs2765997 and rs105334534) of the *PIP4K2A* gene, given that both were more frequent in the Amerindian population, in comparison with the other continental populations, which may reflect a potential correlation with the risk of ALL.

The *IGF2BP1* protein is expressed in a number of different types of cancer, including leukemia. In a recent study based on in vivo and in vitro analyses, Elcheva et al. [32]

found evidence of a significant correlation between *IGF2BP1* and the aggressiveness of leukemia, through the persistence of tumorigenicity by increasing critical transcriptional and metabolic regulators. These authors also emphasized that the *IGF2BP1* gene is often positively regulated in many types of malignant disease, and is not expressed in most normal tissue, which means that it is a potentially important target for anticancer therapy. The results of the present study indicate that the rs8068981 variant of the *IGF2BP1* gene is present at similar frequencies in the Amerindians with ALL and the European population, but contrasts with the Asian populations.

We also identified in the Amerindian population a positive association with the rs3764342 variant of the *WWOX* gene, which has a suppressor role in solid cancers, with the loss of function resulting in alterations in the adhesion of the cancerous cells to the extra-cellular matrix, which affects cell migration and metastasis. A number of studies have contributed to the description of the role of this gene in leukemic malignancies. Luo et al. [33] found that the expression of the *WWOX* mRNA and protein is significantly reduced or absent in cases of leukemia and their cell lines in comparison with the controls, which is consistent with the findings of Cui et al. [34], who evaluated patients diagnosed with different types of leukemia.

In the present study, the rs2304465 variant of the *USP7* gene was also more frequent in Amerindian patients. Jin et al. [35] found an association between *USP7* and a lack of ubiquitination and the stabilization of the *NOTCH1* oncogene, which contributes to the control of the cell growth of the T cells in lymphoblastic leukemia.

5. Conclusions

The present study is the first to investigate the exome of genes involved in the etiology of Acute Lymphoblastic Leukemia (ALL) in Amerindian populations from the Amazon region. The results of the study provide important genetic data related to the etiology of ALL in such population in which genetic investigations are scarce. The study also contributes to the identification of variants of potential risk involved in the etiology of ALL in Amerindian populations, as well as in other Brazilian populations formed through a high level of admixture with this indigenous group.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/jpm12060856/s1>, Table S1. Names and the number of individuals in each population group analyzed in the present study; Table S2. Epidemiological and clinical characteristics of the patients included in the study; Table S3. Description of the variants found in the Amerindian population investigated in the present study; Table S4 [36–48]. Description of the studies selected to choose the genes studied.

Author Contributions: All the authors contributed significantly to the preparation of the manuscript: L.P.C.L. and D.C.d.C. collected and analyzed the data and wrote the text; M.R.F., A.V.W., L.W.M.S.V., N.M.d.S., L.F.P., L.P.A.G., M.C.E.A., S.S.M.F., E.E.B.P., A.M.R.-D.-S. and J.C.G.R. collected and analyzed the data, ran the statistical analyses, and revised the manuscript; J.F.G., Â.R.-d.-S., P.P.d.A., S.E.B.d.S. and N.P.C.d.S. reviewed the statistical analyses, and edited and approved the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (<http://www.cnpq.br>), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (<https://www.gov.br/capes/pt-br>), the Pró-Reitoria de Pesquisa e Pós-Graduação da UFPA (PROPESP) (<http://www.propesp.ufpa.br>), and the Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA) (<http://www.fapespa.pa.gov.br>). This work is part of the Rede de Pesquisa em Genômica Populacional Humana (Biocomputacional-Protocol no. 3381/2013/CAPES).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Brazilian National Commission for Ethics in Research (CONEP) and by the Research Ethics Committee of the UFPA Tropical Medicine Center, under authorization CAAE-20654313.6.0000.5172.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: We acknowledge the Universidade Federal do Pará (UFPA); Oncology Research Center (NPO/UFPA); Postgraduate Program in Genetics and Molecular Biology (PPGBM/UFPA); Human and Medical Genetics Laboratory (LGHM/UFPA); Postgraduate Program in Oncology and Medical Science (PPGOCM/NPO). Ney Santos is supported by CNPq/Produtividade (CNPQ 309999/2021-9).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Pui, C.-H.; Carroll, W.L.; Meshinchi, S.; Arceci, R.J. Biology, Risk Stratification, and Therapy of Pediatric Acute Leukemias: An Update. *J. Clin. Oncol.* **2011**, *29*, 551–565. [[CrossRef](#)] [[PubMed](#)]
2. Bagnasco, F.; Caruso, S.; Andreano, A.; Valsecchi, M.G.; Jankovic, M.; Biondi, A.; Miligi, L.; Casella, C.; Terenziani, M.; Massimino, M.; et al. Late Mortality and Causes of Death among 5-Year Survivors of Childhood Cancer Diagnosed in the Period 1960–1999 and Registered in the Italian Off-Therapy Registry. *Eur. J. Cancer* **2019**, *110*, 86–97. [[CrossRef](#)] [[PubMed](#)]
3. Hunger, S.P.; Mullighan, C.G. Acute Lymphoblastic Leukemia in Children. *N. Engl. J. Med.* **2015**, *373*, 1541–1552. [[CrossRef](#)] [[PubMed](#)]
4. Mullighan, C.G.; Goorha, S.; Radtke, I.; Miller, C.B.; Coustan-Smith, E.; Dalton, J.D.; Girtman, K.; Mathew, S.; Ma, J.; Pounds, S.B.; et al. Genome-Wide Analysis of Genetic Alterations in Acute Lymphoblastic Leukaemia. *Nature* **2007**, *446*, 758–764. [[CrossRef](#)] [[PubMed](#)]
5. Cocco, N.; Anelli, L.; Zagaria, A.; Specchia, G.; Albano, F. Next-Generation Sequencing in Acute Lymphoblastic Leukemia. *Int. J. Mol. Sci.* **2019**, *20*, 2929. [[CrossRef](#)]
6. Papaemmanuil, E.; Hosking, F.J.; Vijaykrishnan, J.; Price, A.; Olver, B.; Sheridan, E.; Kinsey, S.E.; Lightfoot, T.; Roman, E.; Irving, J.A.E.; et al. Loci on 7p12.2, 10q21.2 and 14q11.2 Are Associated with Risk of Childhood Acute Lymphoblastic Leukemia. *Nat. Genet.* **2009**, *41*, 1006–1010. [[CrossRef](#)]
7. Treviño, L.R.; Yang, W.; French, D.; Hunger, S.P.; Carroll, W.L.; Devidas, M.; Willman, C.; Neale, G.; Downing, J.; Raimondi, S.C.; et al. Germline Genomic Variants Associated with Childhood Acute Lymphoblastic Leukemia. *Nat. Genet.* **2009**, *41*, 1001–1005. [[CrossRef](#)]
8. Perez-Andreu, V.; Roberts, K.G.; Harvey, R.C.; Yang, W.; Cheng, C.; Pei, D.; Xu, H.; Gastier-Foster, J.; Shuyu, E.; Lim, J.Y.-S.; et al. Inherited GATA3 Variants Are Associated with Ph-like Childhood Acute Lymphoblastic Leukemia and Risk of Relapse. *Nat. Genet.* **2013**, *45*, 1494–1498. [[CrossRef](#)]
9. Xu, H.; Yang, W.; Perez-Andreu, V.; Devidas, M.; Fan, Y.; Cheng, C.; Pei, D.; Scheet, P.; Burchard, E.G.; Eng, C.; et al. Novel Susceptibility Variants at 10p12.31–12.2 for Childhood Acute Lymphoblastic Leukemia in Ethnically Diverse Populations. *JNCI J. Natl. Cancer Inst.* **2013**, *105*, 733–742. [[CrossRef](#)]
10. Moriyama, T.; Relling, M.V.; Yang, J.J. Inherited Genetic Variation in Childhood Acute Lymphoblastic Leukemia. *Blood* **2015**, *125*, 3988–3995. [[CrossRef](#)]
11. Vijaykrishnan, J.; Studd, J.; Broderick, P.; Kinnersley, B.; Holroyd, A.; Law, P.J.; Kumar, R.; Allan, J.M.; Harrison, C.J.; Moorman, A.V.; et al. Genome-Wide Association Study Identifies Susceptibility Loci for B-Cell Childhood Acute Lymphoblastic Leukemia. *Nat. Commun.* **2018**, *9*, 1340. [[CrossRef](#)] [[PubMed](#)]
12. Yamamoto, J.F.; Goodman, M.T. Patterns of Leukemia Incidence in the United States by Subtype and Demographic Characteristics, 1997–2002. *Cancer Causes Control.* **2008**, *19*, 379–390. [[CrossRef](#)] [[PubMed](#)]
13. Walsh, K.M.; Chokkalingam, A.P.; Hsu, L.-I.; Metayer, C.; de Smith, A.J.; Jacobs, D.I.; Dahl, G.V.; Loh, M.L.; Smirnov, I.V.; Bartley, K.; et al. Associations between Genome-Wide Native American Ancestry, Known Risk Alleles and B-Cell ALL Risk in Hispanic Children. *Leukemia* **2013**, *27*, 2416–2419. [[CrossRef](#)] [[PubMed](#)]
14. Quiroz, E.; Aldoss, I.; Pullarkat, V.; Rego, E.; Marcucci, G.; Douer, D. The Emerging Story of Acute Lymphoblastic Leukemia among the Latin American Population—Biological and Clinical Implications. *Blood Rev.* **2019**, *33*, 98–105. [[CrossRef](#)] [[PubMed](#)]
15. Santos, N.P.C.; Ribeiro-Rodrigues, E.M.; Ribeiro-dos-Santos, A.K.C.; Pereira, R.; Gusmão, L.; Amorim, A.; Guerreiro, J.F.; Zago, M.A.; Matte, C.; Hutz, M.H.; et al. Assessing Individual Interethnic Admixture and Population Substructure Using a 48-Insertion-Deletion (INSEL) Ancestry-Informative Marker (AIM) Panel. *Hum. Mutat.* **2010**, *31*, 184–190. [[CrossRef](#)]
16. Secolin, R.; Mas-Sandoval, A.; Arauna, L.R.; Torres, F.R.; de Araujo, T.K.; Santos, M.L.; Rocha, C.S.; Carvalho, B.S.; Cendes, F.; Lopes-Cendes, I.; et al. Distribution of Local Ancestry and Evidence of Adaptation in Admixed Populations. *Sci. Rep.* **2019**, *9*, 13900. [[CrossRef](#)]
17. Wiemels, J.L.; Walsh, K.M.; de Smith, A.J.; Metayer, C.; Gonseth, S.; Hansen, H.M.; Francis, S.S.; Ojha, J.; Smirnov, I.; Barcellos, L.; et al. GWAS in Childhood Acute Lymphoblastic Leukemia Reveals Novel Genetic Associations at Chromosomes 17q12 and 8q24.21. *Nat. Commun.* **2018**, *9*, 286. [[CrossRef](#)]

18. Ribeiro-Dos-Santos, A.M.; Vidal, A.F.; Vinasco-Sandoval, T.; Guerreiro, J.; Santos, S.; Ribeiro-Dos-Santos, Â.; de Souza, S.J. Exome Sequencing of Native Populations From the Amazon Reveals Patterns on the Peopling of South America. *Front. Genet.* **2020**, *11*, 548507. [[CrossRef](#)]
19. Rodrigues, J.C.G.; de Souza, T.P.; Pastana, L.F.; Ribeiro dos Santos, A.M.; Fernandes, M.R.; Pinto, P.; Wanderley, A.V.; de Souza, S.J.; Kroll, J.E.; Pereira, A.L.; et al. Identification of NUDT15 Gene Variants in Amazonian Amerindians and Admixed Individuals from Northern Brazil. *PLoS ONE* **2020**, *15*, e0231651. [[CrossRef](#)]
20. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff. *Fly* **2012**, *6*, 80–92. [[CrossRef](#)]
21. Shoag, J.M.; Barredo, J.C.; Lossos, I.S.; Pinheiro, P.S. Acute Lymphoblastic Leukemia Mortality in Hispanic Americans. *Leuk. Lymphoma* **2020**, *61*, 2674–2681. [[CrossRef](#)] [[PubMed](#)]
22. Churchman, M.L.; Qian, M.; te Kronnie, G.; Zhang, R.; Yang, W.; Zhang, H.; Lana, T.; Tedrick, P.; Baskin, R.; Verbist, K.; et al. Germline Genetic IKZF1 Variation and Predisposition to Childhood Acute Lymphoblastic Leukemia. *Cancer Cell* **2018**, *33*, 937–948.e8. [[CrossRef](#)] [[PubMed](#)]
23. Hong, Y.; Wang, Q.; Song, Y.L.; Liu, G.M.; Yang, X.; Yan, M.X. Relationship between IKZF3 Gene Single Nucleotide Polymorphisms and Childhood Acute Lymphoblastic Leukemia. *Zhongguo Shi Yan Xue Ye Xue Za Zhi* **2021**, *29*, 690–695. [[CrossRef](#)] [[PubMed](#)]
24. Shi, Y.; Du, M.; Fang, Y.; Tong, N.; Zhai, X.; Sheng, X.; Li, Z.; Xue, Y.; Li, J.; Chu, H.; et al. Identification of a Novel Susceptibility Locus at 16q23.1 Associated with Childhood Acute Lymphoblastic Leukemia in Han Chinese. *Hum. Mol. Genet.* **2016**, *25*, 2873–2880. [[CrossRef](#)] [[PubMed](#)]
25. Dong, Y.; Wang, X.; Yang, Y.-W.; Liu, Y.-J. The Effects of CDKN2A Rs3731249, Rs11515, and Rs3088440 Polymorphisms on Cancer Risk. *Cell. Mol. Biol.* **2017**, *63*, 40. [[CrossRef](#)] [[PubMed](#)]
26. Royds, J.A.; Pilbrow, A.P.; Ahn, A.; Morrin, H.R.; Frampton, C.; Russell, I.A.; Moravec, C.S.; Sweet, W.E.; Tang, W.H.W.; Currie, M.J.; et al. The Rs11515 Polymorphism Is More Frequent and Associated with Aggressive Breast Tumors with Increased ANRIL and Decreased P16INK4a Expression. *Front. Oncol.* **2016**, *5*, 306. [[CrossRef](#)]
27. Sibin, M.K.; Dhananjaya, I.B.; Narasingarao, K.V.L.; Harshitha, S.M.; Jeru-Manoj, M.; Chetan, G.K. Two Gene Polymorphisms (Rs4977756 and Rs11515) in CDKN2A/B and Glioma Risk in South Indian Population. *Meta Gene* **2016**, *9*, 215–218. [[CrossRef](#)]
28. Sauroja, I.; Smeds, J.; Vlaykova, T.; Kumar, R.; Talve, L.; Hahka-Kemppinen, M.; Punnonen, K.; Jansèn, C.T.; Hemminki, K.; Pyrhönen, S. Analysis of G1/S Checkpoint Regulators in Metastatic Melanoma. *Genes Chromosomes Cancer* **2000**, *28*, 404–414. [[CrossRef](#)]
29. McCloud, J.; Sivakumar, R.; Greenhough, A.; Elder, J.; Jones, P.; Deakin, M.; Elder, J.; Fryer, A.; Hoban, P. P16INK4a Polymorphism: Associations with Tumour Progression in Patients with Sporadic Colorectal Cancer. *Int. J. Oncol.* **2004**, *25*, 1447–1452. [[CrossRef](#)]
30. Liao, F.; Yin, D.; Zhang, Y.; Hou, Q.; Zheng, Z.; Yang, L.; Shu, Y.; Xu, H.; Li, Y. Association Between PIP4K2A Polymorphisms and Acute Lymphoblastic Leukemia Susceptibility. *Medicine* **2016**, *95*, e3542. [[CrossRef](#)]
31. Zhou, X.; Liao, F.; Zhang, J.; Qin, Y.; Xu, H.; Ding, Z.; Zhang, Y.; Zhang, F. Association of the Independent Polymorphisms in CDKN2A with Susceptibility of Acute Lymphoblastic Leukemia. *Biosci. Rep.* **2018**, *38*. [[CrossRef](#)] [[PubMed](#)]
32. Elcheva, I.A.; Wood, T.; Chiarolanzio, K.; Chim, B.; Wong, M.; Singh, V.; Gowda, C.P.; Lu, Q.; Hafner, M.; Dovat, S.; et al. RNA-Binding Protein IGF2BP1 Maintains Leukemia Stem Cell Properties by Regulating HOXB4, MYB, and ALDH1A1. *Leukemia* **2020**, *34*, 1354–1363. [[CrossRef](#)] [[PubMed](#)]
33. Luo, L.; Chen, Y.; Cheng, X.; Lin, Y.; Fu, X.; Li, D.; Cui, Z.; Lin, D. Reduced Expression of the WW Domain-Containing Oxidoreductase in Human Hematopoietic Malignancies. *Oncol. Lett.* **2016**, *11*, 4083–4088. [[CrossRef](#)]
34. Cui, Z.; Lin, D.; Cheng, F.; Luo, L.; Kong, L.; Xu, J.; Hu, J.; Lan, F. The Role of the WWOX Gene in Leukemia and Its Mechanisms of Action. *Oncol. Rep.* **2013**, *29*, 2154–2162. [[CrossRef](#)]
35. Jin, Q.; Martinez, C.A.; Arcipowski, K.M.; Zhu, Y.; Gutierrez-Diaz, B.T.; Wang, K.K.; Johnson, M.R.; Volk, A.G.; Wang, F.; Wu, J.; et al. USP7 Cooperates with NOTCH1 to Drive the Oncogenic Transcriptional Program in T-Cell Leukemia. *Clin. Cancer Res.* **2019**, *25*, 222–239. [[CrossRef](#)] [[PubMed](#)]
36. Hao, Q.; Cao, M.; Zhang, C.; Yin, D.; Wang, Y.; Ye, Y.; Zhao, S.; Yang, Y.; Chen, K.-L.; Ying, B.; et al. Age-related differences of genetic susceptibility to patients with acute lymphoblastic leukemia. *Aging* **2021**, *13*, 12456–12465. [[CrossRef](#)] [[PubMed](#)]
37. Perez-Andreu, V.; Roberts, K.G.; Xu, H.; Smith, C.; Zhang, H.; Yang, W.; Harvey, R.C.; Payne-Turner, D.; Devidas, M.; Cheng, I.-M.; et al. A genome-wide association study of susceptibility to acute lymphoblastic leukemia in adolescents and young adults. *Blood* **2015**, *125*, 680–686. [[CrossRef](#)]
38. Urayama, K.Y.; Takagi, M.; Kawaguchi, T.; Matsuo, K.; Tanaka, Y.; Ayukawa, Y.; Arakawa, Y.; Hasegawa, D.; Yuza, Y.; Kaneko, T.; et al. Regional evaluation of childhood acute lymphoblastic leukemia genetic susceptibility loci among Japanese. *Sci. Rep.* **2018**, *8*, 789. [[CrossRef](#)]
39. Semmes, E.C.; Vijayakrishnan, J.; Zhang, C.; Hurst, J.H.; Houlston, R.S.; Walsh, K.M. Leveraging Genome and Phenome-Wide Association Studies to Investigate Genetic Risk of Acute Lymphoblastic Leukemia. *Cancer Epidemiol. Biomarkers Prev.* **2020**, *29*, 1606–1614. [[CrossRef](#)]
40. Vijayakrishnan, J.; Qian, M.; Studd, J.B.; Yang, W.; Kinnersley, B.; Law, P.J.; Broderick, P.; Raetz, E.A.; Allan, J.; Pui, C.-H.; et al. Identification of four novel associations for B-cell acute lymphoblastic leukaemia risk. *Nat. Commun.* **2019**, *10*, 1–9. [[CrossRef](#)]

41. Brown, A.L.; de Smith, A.J.; Gant, V.U.; Yang, W.; Scheurer, M.E.; Walsh, K.M.; Chernus, J.M.; Kallsen, N.A.; Peyton, S.A.; Davies, G.E.; et al. Inherited genetic susceptibility to acute lymphoblastic leukemia in Down syndrome. *Blood* **2019**, *134*, 1227–1237. [[CrossRef](#)] [[PubMed](#)]
42. Sherborne, A.; Hosking, F.; Prasad, R.; Kumar, R.; Koehler, R.; Vijayakrishnan, J.; Papaemmanuil, E.; Bartram, C.R.; Stanulla, M.; Schrappe, M.; et al. Variation in CDKN2A at 9p21.3 influences childhood acute lymphoblastic leukemia risk. *Nat. Genet.* **2010**, *42*, 492–494. [[CrossRef](#)]
43. Hungate, E.A.; Vora, S.R.; Gamazon, E.; Moriyama, T.; Best, T.; Hular, I.; Lee, Y.; Evans, T.-J.; Ellinghaus, E.; Stanulla, M.; et al. A variant at 9p21.3 functionally implicates CDKN2B in paediatric B-cell precursor acute lymphoblastic leukaemia aetiology. *Nat. Commun.* **2016**, *7*, 10635. [[CrossRef](#)]
44. Liao, F.; Ye, Y.; Yin, D.; Qin, Y.; Zhao, J.; Zhang, W.; Zhang, Y.; Deng, Z.; Wang, Y.; Ying, B.; et al. Validations of Top and Novel Susceptibility Variants in All-Age Chinese Patients With Acute Lymphoblastic Leukemia. *Front. Genet.* **2020**, *11*, 1004. [[CrossRef](#)] [[PubMed](#)]
45. Al-Absi, B.; Razif, M.F.M.M.F.; Noor, S.M.S.M.; Saif-Ali, R.; Aqlan, M.; Salem, S.D.S.D.; Ahmed, R.H.R.H.; Muniandy, S. Contributions of IKZF1, DDC, CDKN2A, CEBPE, and LMO1 Gene Polymorphisms to Acute Lymphoblastic Leukemia in a Yemeni Population. *Genet. Test. Mol. Biomark.* **2017**, *21*, 592–599. [[CrossRef](#)] [[PubMed](#)]
46. Vijayakrishnan, J.; Kumar, R.; Henrion, M.Y.R.; Moorman, A.V.; Rachakonda, P.S.; Hosen, I.; Filho, M.I.D.S.; Holroyd, A.; E Dobbins, S.; Koehler, R.; et al. A genome-wide association study identifies risk loci for childhood acute lymphoblastic leukemia at 10q26.13 and 12q23.1. *Leukemia* **2016**, *31*, 573–579. [[CrossRef](#)]
47. Zhang, H.; Liu, A.P.-Y.; Devidas, M.; Lee, S.H.; Cao, X.; Pei, D.; Borowitz, M.; Wood, B.; Gastier-Foster, J.M.; Dai, Y.; et al. Association of GATA3 Polymorphisms With Minimal Residual Disease and Relapse Risk in Childhood Acute Lymphoblastic Leukemia. *JNCI: J. Natl. Cancer Inst.* **2020**, *113*, 408–417. [[CrossRef](#)]
48. Qian, M.; Zhao, X.; Devidas, M.; Yang, W.; Gocho, Y.; Smith, C.; Gastier-Foster, J.M.; Li, Y.; Xu, H.; Zhang, S.; et al. Genome-Wide Association Study of Susceptibility Loci for T-Cell Acute Lymphoblastic Leukemia in Children. *JNCI: J. Natl. Cancer Inst.* **2019**, *111*, 1350–1357. [[CrossRef](#)]