Published in final edited form as:
Leuk Res. 2013 August ; 37(8): 936–942. doi:10.1016/j.leukres.2013.04.022.

ARID5B and *IKZF1* variants, selected demographic factors, and childhood acute lymphoblastic leukemia: A report from the Children's Oncology Group

Amy M. Linabery, PhD[1], **Crystal N. Blommer**[2], **Logan G. Spector, PhD**[1,2], **Stella M. Davies, MBBS, PhD**[3], **Leslie L. Robison, PhD**[4], and **Julie A. Ross, PhD**[1,2]

[1]Division of Pediatric Epidemiology and Clinical Research, Department of Pediatrics, University of Minnesota, Minneapolis, MN, USA

[2]University of Minnesota Masonic Cancer Center, Minneapolis, MN, USA

[3]Division of Hematology/Oncology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

[4]Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, TN, USA

Abstract

Interactions between common germline variants in *ARID5B* and *IKZF1* and other known childhood acute lymphoblastic leukemia (ALL) risk factors were queried using biospecimens and data from 770 ALL cases and 384 controls. Case-control comparisons revealed dosedependent associations between *ARID5B* rs10821936, *ARID5B* rs10994982, and *IKZF1* rs11978267 and childhood ALL overall, and B lineage and B lineage hyperdiploid ALL examined separately (all allelic odds ratios≥1.33, P_{trend}≤0.001). No heterogeneity was observed between ORs for males and females (all $P_{interaction}$≥0.48). Likewise, no significant genotype-birth weight interactions were detected (all $P_{interaction}$≥0.12) among cases. These results indicate similar ALL risk across strata of known risk factors.

Keywords

acute lymphoblastic leukemia; children; genetic susceptibility; gene-environment interaction

To whom correspondence should be addressed: Amy Linabery, PhD; Department of Pediatrics; University of Minnesota; 420 Delaware Street SE, MMC 715; Minneapolis, MN 55455; Telephone: 1-612-626-0278, Fax: 1-612-624-7147; linabery@umn.edu.

Authors' Contributions

All authors made substantial contributions to this manuscript. A.M.L. participated in study development and implementation, was responsible for the statistical analysis, and was primary author of all sections of the manuscript. C.N.B. conducted laboratory experiments on the biospecimens and contributed to the final manuscript. L.G.S. and S.M.D. each contributed meaningfully to interpreting study results and to editing the final manuscript. L.L.R. was principal investigator of the original CCG-E15 case-control study and thereby contributed substantially to sample and data collection; he also participated in data interpretation and revision of the final manuscript. J.A.R. was responsible for the conception and implementation of the study, supervised laboratory analyses, and played a significant role in data interpretation and in writing up the study results.

Conflict of Interest

The authors declare no conflict of interest.

Introduction

Few risk factors have been identified for childhood ALL, the most common pediatric malignancy occurring in 38/1 000 000 children in the United States per year.[1] In addition to rare genetic syndromes, which account for ~5% or less of childhood ALLs, established factors conferring increased risk include male sex, white race, Hispanic ethnicity, and *in utero* exposure to irradiation.[2,3] High birth weight and advanced maternal age have also been consistently and positively associated with childhood ALL.[4,5] Further, recent studies have identified common germline variants contributing genetic susceptibility for childhood ALL.

Two independent genome-wide association studies (GWAS) of pediatric ALL revealed consistent significant associations with single nucleotide polymorphisms (SNPs) at chromosome 10 bands q11.22 and q21.2, the location of *ARID5B* (AT-rich active interactive domain 5B), and at chromosome 7 band p12.2, the site of *IKZF1* (Ikaros family zinc finger 1).[6,7] Importantly, the loci showed specificity for precursor B cell ALL and for the B-hyperdiploid subtype,[6,7] and were confirmed to be independently associated with childhood B cell ALL in subsequent GWAS and validation studies.[8–16] Germline variants in these genes have also been associated with ALL in infants,[17] but not in adults,[18] and with AML in infants[17] and children.[19] These observations are compelling, given that the transcription factors encoded by these genes are involved in regulating lymphocyte differentiation.[20–22] Somatic mutations have also been reported: *IKZF1* is frequently mutated or deleted in ALL subtypes associated with poorer prognosis, including *BCR-ABL1*-positive ALL and ALL with a *BCR-ABL1*-like expression pattern,[23,24] while deletions in *IKZF1* and *ARID5B* have been reported in those with B hyperdiploid ALL.[25]

Little is currently known about the joint distributions of these SNPs and other known or suspected risk factors. Although reported variant allele frequencies have varied noticeably by case race and ethnicity, resulting associations have been comparable across racial and ethnic groups due to corresponding allele frequency differences in the underlying control populations.[10,11,14] *ARID5B* genotype distributions were also similar between those diagnosed at ≥10 versus <10 years of age in one large study,[14] while allele frequencies or associations differed significantly between males and females for some *ARID5B* SNPs, but not others.[9,14] Overall, prior research suggests the causal risk variants annotated by these SNPs are generalizable to B lineage ALL across populations and demographic groups.

The purpose of the current analysis was to investigate possible risk differences across strata of confirmed risk factors using interview and clinical data from a North American case series of pediatric ALL.

Materials and Methods

Childhood ALL cases

Cases consisted of patients enrolled in Children's Cancer Group (CCG) (now Children's Oncology Group (COG)) Protocol E15 – Epidemiology of ALL Study with biospecimens available for genotyping.[26,27] These patients included children with a confirmed diagnosis of ALL at ages 0–14 years at one of the 108 participating U.S. CCG institutions during the period January 1, 1989-June 15, 1993. After excluding 5 patients with Down syndrome, 839 ALL cases were eligible. Of these, 69 (8.2%) could not be genotyped for any of the 3 selected SNPs, leaving 770 cases for analysis. Distributions of diagnostic age, sex, birth weight, and maternal age were similar between the 770 included and 1144 excluded cases from the original CCG-E15 case-control study (all $P \geq 0.29$); however, there were more

included cases with known cell lineage (87% vs. 82%, P=0.0003) and more non-Hispanic white (84% vs. 77%) and fewer Hispanic cases (6% vs. 11%, P=0.002).

Data on demographic and birth characteristics (sex, race/ethnicity, birth weight, maternal age at child's birth), and other exposures, were collected via maternal interview.[26,27] Diagnostic information and archived or cryopreserved bone marrow samples were also collected.[26,27] Bone marrow specimens were immunophenotyped for B or T cell lineage via a standard panel of monoclonal antibodies; a subset were unclassifiable.[26] Chromosomal numerical abnormalities (ploidy), including hyperdiploidy (defined as >50 chromosomes per leukemia cell), was also evaluated in a subset of cases. DNA was isolated from archived biological samples using a phenol-chloroform extraction protocol.[27,28] All isolated DNA was stored at −80°C prior to genotyping.

Control samples

No biospecimens were obtained from controls in the original CCG-E15 study, so anonymized blood samples from 384 healthy, non-Hispanic white blood donors (207 males, 177 females) were used as controls. DNA was isolated via a DNA blood midi column protocol (Qiagen, Inc., Valencia, CA) and was stored at −20°C prior to genotyping.

Genotyping methods

Selected SNPs (*ARID5B* rs10821936 and rs10994982, and *IKZF1* rs11978267) were genotyped by fluorogenic PCR-based allelic discrimination (Taqman) assays using the ABI Prism 7900HT RT-PCR System (Applied Biosystems, Inc., Carlsbad, CA). Primers and probes were designed by ABI (Taqman SNP assays C_26140184_10, C_30824850_10, and C_199413_10). PCR conditions are available on request. Genotype calls were made manually upon visual inspection of allelic discrimination plots by identifying clusters around reference controls for each allele.

Genotypes could not be called for 0.4–7% of the 770 case samples across the 3 SNPs. For the control samples, the call rate was ≥99.7% for each SNP. As a quality control measure, ~6% of case and control samples were selected for duplication for each SNP, resulting in 100% concordance between replicates. In addition, no deviations from Hardy-Weinberg Equilibrium (HWE), evaluated via Pearson's chi-square test, were observed in control samples for any of the loci (all P≥0.65), providing additional confirmation of genotyping accuracy.

The University of Minnesota Institutional Review Board approved the current study; IRBs of participating CCG institutions previously approved the parent CCG-E15 case-control study.

Statistical analysis

Case genotypes were compared to those of controls at each locus via unconditional logistic regression. Genomic odds ratios (ORs) and 95% confidence intervals (CIs) were calculated to evaluate individual genotypes (AB vs. AA, BB vs. AA) and allelic ORs were estimated from ordinal variables indicating the number of high risk alleles (0, 1, 2) to assess the risk per additional copy of the variant allele (dose response). The ORs are unadjusted, as little covariate data were available on controls and the inclusion of sex in the allelic models did not change the ln(OR) estimates appreciably. P-values for trend were determined from allelic models via the Cochran-Armitage trend test in all groups except the age subgroups, where the Wald chi-square test was used. Epistasis was evaluated by the inclusion of pairwise interaction terms in logistic regression models where genotypes were modeled as

ordinal variables. Sex-specific case-control comparisons were also conducted and genotype-sex interactions were tested in the allelic models.

Case-only models were used to examine potential gene-by-environment interactions on the multiplicative scale using ordinal logistic regression, where the independent variables were the demographic characteristics (categories of birth weight and maternal age) and the dependent variables were the genotypes modeled as the number of high risk alleles (0, 1, 2). The resulting allelic interaction odds ratios (IORs) were presented in instances where the proportional hazards assumption was not violated. Alternatively, genomic IORs were provided in table footnotes for models where the proportional hazards assumption did not hold. Minimal detectable IORs were estimated using Quanto version 1.2.4,[29] given the sample size of 770 and allele frequencies of 0.36–0.57, and assuming population prevalences of 7.2% for birth weight >4000 g and 12.5% for maternal age >35 years (from E15 controls), a log-additive mode of inheritance, a type I error rate of 0.05, and 80% power.

Polytomous logistic regression was performed to calculate and test the homogeneity of the ORs for the diagnosis age subgroups compared with controls. Due to the potential for confounding by population admixture and given the known population differences in variant allele frequencies for these SNPs,[10,11,14] all analyses were repeated restricting cases to non-Hispanic white (n=645) by maternal report.

All regression analyses were conducted in SAS version 9.2 (SAS Institute Inc., Cary, NC, USA).

Results

Characteristics of the 770 cases are shown in Table 1. Most ALLs were of a B cell lineage (74.5%). Over half were diagnosed between 2–5 years of age (54.8%) and a similar proportion were male (55.1%). Distributions of ALL subtype by diagnostic age are provided in Supplementary Table 1.

Risk allele frequencies were greater for cases than controls for the *ARID5B* rs10821936 C (0.44 and 0.34, respectively, Table 2), *ARID5B* rs10994982 A (0.57 and 0.49), and *IKZF1* rs11978267 G (0.36 and 0.27) variants and were similar to previously published frequencies.[6] These differences correspond to statistically significant associations between additional copies of the variants, indicating a positive dose response, and childhood ALL overall (allelic $OR_{rs10821936}$=1.56, 95% CI: 1.31–1.87, $P_{trend}=9\times10^{-7}$; $OR_{rs10994982}$=1.33, 95% CI: 1.12–1.58, P_{trend}=0.001; $OR_{rs11978267}$=1.45, 95% CI: 1.20–1.74, P_{trend}=0.0001; Table 2), as well as for B lineage ALL and for B lineage hyperdiploid ALL (all allelic OR≥1.37, $P_{trend}\leq0.001$; Table 2). None of the three SNPs was associated with T-cell ALL. In evaluating potential epistasis, there were no pairwise interactions detected between any of the three SNPs (all $P_{interaction}\geq0.24$; data not shown) for childhood ALL overall or among the ALL subtypes.

The SNP associations were similar for males and females for each of the three SNPs overall and for each of the subtypes (all $P_{interaction}\geq0.48$, Table 3). There were no interactions detected between any of the three SNPs and birth weight (all $P_{interaction}\geq0.12$) (Table 4). For maternal age, a significant interaction was observed in the genomic model for *ARID5B* rs10994982 due to a lower frequency of the A allele in cases of mothers aged 20–24 years ($P_{interaction}$=0.02). SNP associations were restricted to the intermediate diagnostic age categories spanning 1–10 years and were not observed in infants or in those aged 11–14 years (all $P_{heterogeneity}\leq0.02$, Supplementary Table 2). Restricting to non-Hispanic white cases only (n=645) produced results very similar to those described above (data not shown).

Discussion

Here we further validate associations between childhood ALL and *ARID5B* and *IKZF1* SNPs identified through GWAS;[6,7] we report unusually strong associations for B lineage hyperdiploid ALL, but none for T cell ALL. Importantly, we report no marked differences in the associations across strata of sex, birth weight, or maternal age at the time of the child's birth. The lack of detectable epistasis suggests that each risk variant contributes independently and additively to the risk of childhood B cell ALL, perhaps through altered transcriptional regulation of lymphocyte differentiation.

Ikaros and ARID5B participate as transcription factors in the regulation of normal immune cell maturation. The effects of Ikaros begin early in hematopoiesis, assisting hematopoietic stem cell maintenance and renewal[30] and impacting lineage commitment decisions and differentiation of lymphoid cell populations.[20] With respect to early B cell lymphopoiesis, Ikaros is instrumental in regulating genes involved in expression of cell surface receptors and V(D)J recombination [21]. Using Epstein-Barr virus-transformed lymphocytes, Papaemmanuil *et al* showed a functional effect for rs4132601,[7] which is in complete linkage disequilibrium with rs11978267,[31] such that each additional copy of the variant allele resulted in greater attenuation of *IKZF1* mRNA expression. Additionally, substantial attenuation of *Ikzf1* expression in *BCR-ABL1*$^{+/0}$ mice heterozygous for a Lac Z knock-in in exon 2 (IK$^{L/+}$) resulted in an increased population of lymphocytes arrested at the pro-B to pre-B stage, which quickly progressed to leukemia in all mice tested.[32] This and similar evidence from other murine models of *Ikzf1* loss-of-function mutations,[33] along with genomic profiling studies of human leukemia samples showing a high prevalence of *IKZF1* mutations and deletions leading to haploinsufficiency,[23–25,34] provides robust support for a potent tumor suppressor role for Ikaros in pediatric B cell ALL. The increased susceptibility for B lineage ALL observed here and by others[6–16] suggests that rs11978267 is associated with ALL evolving in B cell progenitors in the early stages of differentiation.

The role of ARID5B in lymphocyte differentiation and function has been less thoroughly investigated. A pivotal piece of evidence is that loss of *Arid5b* expression in mice with homozygous deletions within the conserved ARID region results in transient reductions in total immune cell counts in the bone marrow and secondary lymphoid organs, as well as reduced proportions of lymphocyte progenitors in those organs.[22] Although the functional consequences of the two *ARID5B* SNPs examined herein are not known, it is possible that the presence of a variant allele may also lead to transient abnormalities in immune cell distributions and increased susceptibility to leukemia.

Associations between the variant alleles and ALL were similar across both sexes. A greater male:female ratio is observed for T cell ALL (2.4:1);[35] however, since none of the SNPs were associated with T cell ALL here, or in most prior studies, this homogeneity of ORs would be expected. These sex-specific results replicate those reported by Xu *et al* and Healy *et al* for *ARID5B* rs10821936,[9,14] but for *ARID5B* rs10994982 our results contradict those by Healy *et al*, who observed a 3.8-fold greater odds of ALL in males homozygous for the variant compared to wild type homozygotes, but no association in females (P=0.01).[9] In contrast, Orsi *et al* reported ORs that were 30–40% higher among females for *ARID5B* SNPs in high linkage disequilibrium with those listed above (interaction P-values were not provided); however, similar to our results, they found comparable sex-specific ORs for the *IKZF1* locus.[16]

We did not observe significant differences in the distribution of genotypes across stratum of birth weight, or maternal age for 2 of 3 SNPs, although we had 80% power to detect IORs of

~1.5 for the highest categories of exposures (e.g., birth weights >4000g) in the case-only models. For maternal age, a statistical interaction was detected for *ARID5B* rs10994982, where the risk allele frequency was somewhat lower in cases born to mothers aged 20–24 years than in other categories; however, this may be a chance finding because there was no discernable pattern of risk across maternal age categories.

When risk differences were stratified by age at diagnosis, associations were not observed among infant or adolescent cases, although these groups have smaller sample sizes and results should be interpreted accordingly. As described in our similar study of infant leukemia,[17] these null findings for infants likely reflect an absence of an association within infant ALL cases with *MLL* gene rearrangements *(MLL+),* for whom an alternate pathway is likely involved. The current study preceded routine clinical *MLL* testing, however, it is likely that a substantial portion of the infants were ALL *MLL+*, since an estimated 61–79% of infant ALLs fall into this category.[36,37] With respect to adolescents, the greater proportion of cases with T lineage ALL (33%) resulted in no SNP associations in that age group.

The present study has several strengths, including the sample size and the inclusion of case samples and data from a North American epidemiologic study. *Post hoc* power calculations[29] in the combined ALL cases and in B lineage ALLs indicated there was 89–100% power to detect the observed associations. For T lineage ALLs, there was sufficient power (≥80%) to detect ORs ≥1.6, had such effects been present. In restricting our analysis of gene-environment interactions to those SNPs consistently associated in prior GWAS with childhood ALL and factors that are both reliably recalled[38] and consistently associated with ALL,[4,5] we minimize the potential for recall bias and chance findings. The main limitation is that we are using data from a historical ALL study that pre-dates routine cytogenetic analysis for recurring translocations (e.g., *ETV6-RUNX1*[15], *BCR-ABL1*), and for whom karyotypes are no longer available, precluding our ability to further stratify by these biologically meaningful characteristics. In addition, a number of statistical tests were performed in the current study. We did not adjust for multiple comparisons in order to detect associations, however we acknowledge that ORs with *P*-values near the 0.05 significance level may be spurious. Finally, we have focused on three SNPs in two genes, however it is not yet clear if these polymorphisms are causal variants or markers of causal variants, and it is further possible that we have missed etiologically relevant variants through this approach.

Conclusions

In a North American case series we have shown that the confirmed associations between variant alleles in *ARID5B* and *IKZF1* and B lineage childhood ALL do not vary by sex, birth weight, or maternal age categories. Associations did differ across the age at diagnosis categories, however, which reflects the distribution of ALL subtypes present in each category.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health [grant numbers K05 CA157439, U10CA13539, U10CA98543]; and the Children's Cancer Research Fund, Minneapolis, MN.

References

1. Linabery AM, Ross JA. Trends in childhood cancer incidence in the U.S (1992–2004). Cancer. 2008; 112:416–32. [PubMed: 18074355]
2. Pizzo, PA.; Poplack, DG. Principles and practice of pediatric oncology. 6. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2011.
3. Wakeford R, Little MP. Risk coefficients for childhood cancer after intrauterine irradiation: a review. Int J Radiat Biol. 2003; 79:293–309. [PubMed: 12943238]
4. Caughey RW, Michels KB. Birth weight and childhood leukemia: a meta-analysis and review of the current evidence. Int J Cancer. 2009; 124:2658–70. [PubMed: 19173295]
5. Johnson KJ, Carozza SE, Chow EJ, Fox EE, Horel S, McLaughlin CC, et al. Parental age and risk of childhood cancer: a pooled analysis. Epidemiology. 2009; 20:475–83. [PubMed: 19373093]
6. Trevino LR, Yang W, French D, Hunger SP, Carroll WL, Devidas M, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nature Genetics. 2009; 41:1001–5. [PubMed: 19684603]
7. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, et al. Loci on 7p12.2, 10q21.2 and 14q11. 2 are associated with risk of childhood acute lymphoblastic leukemia. Nature Genetics. 2009; 41:1006–10. [PubMed: 19684604]
8. Prasad RB, Hosking FJ, Vijayakrishnan J, Papaemmanuil E, Koehler R, Greaves M, et al. Verification of the susceptibility loci on 7p12.2, 10q21.2, and 14q11. 2 in precursor B-cell acute lymphoblastic leukemia of childhood. Blood. 2010; 115:1765–7. [PubMed: 20042726]
9. Healy J, Richer C, Bourgey M, Kritikou EA, Sinnett D. Replication analysis confirms the association of ARID5B with childhood B-cell acute lymphoblastic leukemia. Haematologica. 2010; 95:1608–11. [PubMed: 20460642]
10. Yang W, Trevino LR, Yang JJ, Scheet P, Pui CH, Evans WE, et al. ARID5B SNP rs10821936 is associated with risk of childhood acute lymphoblastic leukemia in blacks and contributes to racial differences in leukemia incidence. Leukemia. 2010; 24:894–6. [PubMed: 20054350]
11. Vijayakrishnan J, Sherborne AL, Sawangpanich R, Hongeng S, Houlston RS, Pakakasama S. Variation at 7p12.2 and 10q21. 2 influences childhood acute lymphoblastic leukemia risk in the Thai population and may contribute to racial differences in leukemia incidence. Leuk Lymphoma. 2010; 51:1870–4. [PubMed: 20919861]
12. Han S, Lee KM, Park SK, Lee JE, Ahn HS, Shin HY, et al. Genome-wide association study of childhood acute lymphoblastic leukemia in Korea. Leuk Res. 2010; 34:1271–4. [PubMed: 20189245]
13. Pastorczak A, Gorniak P, Sherborne A, Hosking F, Trelinska J, Lejman M, et al. Role of 657del5 NBN mutation and 7p12.2 (IKZF1), 9p21 (CDKN2A), 10q21.2 (ARID5B) and 14q11. 2 (CEBPE) variation and risk of childhood ALL in the Polish population. Leuk Res. 2011; 35:1534–6. [PubMed: 21889209]
14. Xu H, Cheng C, Devidas M, Pei D, Fan Y, Yang W, et al. ARID5B genetic polymorphisms contribute to racial disparities in the incidence and treatment outcome of childhood acute lymphoblastic leukemia. J Clin Oncol. 2012; 30:751–7. [PubMed: 22291082]
15. Ellinghaus E, Stanulla M, Richter G, Ellinghaus D, Te Kronnie G, Cario G, et al. Identification of germline susceptibility loci in ETV6-RUNX1-rearranged childhood acute lymphoblastic leukemia. Leukemia. 2012; 26:902–9. [PubMed: 22076464]
16. Orsi L, Rudant J, Bonaventure A, Goujon-Bellec S, Corda E, Evans TJ, et al. Genetic polymorphisms and childhood acute lymphoblastic leukemia: GWAS of the ESCALE study (SFCE). Leukemia. 2012; 26:2561–4. [PubMed: 22660188]
17. Ross JA, Linabery AM, Blommer CN, Langer EK, Spector LG, Hilden JM, et al. Genetic variants modify susceptibility to leukemia in infants: A Children's Oncology Group report. Pediatric Blood Cancer. 2013; 60:31–4. [PubMed: 22422485]
18. Peyrouze P, Guihard S, Grardel N, Berthon C, Pottier N, Pigneux A, et al. Genetic polymorphisms in ARID5B, CEBPE, IKZF1 and CDKN2A in relation with risk of acute lymphoblastic leukaemia in adults: a Group for Research on Adult Acute Lymphoblastic Leukaemia (GRAALL) study. Br J Haematol. 2012; 159:599–602. [PubMed: 23016962]

19. Rudant J, Orsi L, Bonaventure A, Goujon-Bellec S, Corda E, Baruchel A, et al. Are ARID5B and IKZF1 polymorphisms also associated with childhood acute myeloblastic leukemia: the ESCALE study (SFCE)? Leukemia. 2012 [Epub ahead of print]. 10.1038/leu.2012.244
20. Georgopoulos K, Bigby M, Wang JH, Molnar A, Wu P, Winandy S, et al. The Ikaros gene is required for the development of all lymphoid lineages. Cell. 1994; 79:143–56. [PubMed: 7923373]
21. Sellars M, Kastner P, Chan S. Ikaros in B cell development and function. World J Biol Chem. 2011; 2:132–9. [PubMed: 21765979]
22. Lahoud MH, Ristevski S, Venter DJ, Jermiin LS, Bertoncello I, Zavarsek S, et al. Gene targeting of Desrt, a novel ARID class DNA-binding protein, causes growth retardation and abnormal development of reproductive organs. Genome Res. 2001; 11:1327–34. [PubMed: 11483573]
23. Mullighan CG, Miller CB, Radtke I, Phillips LA, Dalton J, Ma J, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. Nature. 2008; 453:110–4. [PubMed: 18408710]
24. Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med. 2009; 360:470–80. [PubMed: 19129520]
25. Paulsson K, Forestier E, Lilljebjorn H, Heldrup J, Behrendtz M, Young BD, et al. Genetic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. Proc Natl Acad Sci USA. 2010; 107:21719–24. [PubMed: 21098271]
26. Brondum J, Shu XO, Steinbuch M, Severson RK, Potter JD, Robison LL. Parental cigarette smoking and the risk of acute leukemia in children. Cancer. 1999; 85:1380–8. [PubMed: 10189146]
27. Shu XO, Perentesis JP, Wen W, Buckley JD, Boyle E, Ross JA, et al. Parental exposure to medications and hydrocarbons and ras mutations in children with acute lymphoblastic leukemia: a report from the Children's Oncology Group. Cancer Epidemiol Biomarkers Prev. 2004; 13:1230–5. [PubMed: 15247135]
28. Boyle EB, Steinbuch M, Tekautz T, Gutman JR, Robison LL, Perentesis JP. Accuracy of DNA amplification from archival hematological slides for use in genetic biomarker studies. Cancer Epidemiol Biomarkers Prev. 1998; 7:1127–31. [PubMed: 9865432]
29. Gauderman, WJ.; Morrison, JM. QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. 2006. http://hydra.usc.edu/gxe
30. Papathanasiou P, Attema JL, Karsunky H, Hosen N, Sontani Y, Hoyne GF, et al. Selfrenewal of the long-term reconstituting subset of hematopoietic stem cells is regulated by Ikaros. Stem Cells. 2009; 27:3082–92. [PubMed: 19816952]
31. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008; 24:2938–9. [PubMed: 18974171]
32. Virely C, Moulin S, Cobaleda C, Lasgi C, Alberdi A, Soulier J, et al. Haploinsufficiency of the IKZF1 (IKAROS) tumor suppressor gene cooperates with BCR-ABL in a transgenic model of acute lymphoblastic leukemia. Leukemia. 2010; 24:1200–4. [PubMed: 20393504]
33. Kastner P, Chan S. Role of Ikaros in T-cell acute lymphoblastic leukemia. World J Biol Chem. 2011; 2:108–14. [PubMed: 21765975]
34. Payne KJ, Dovat S. Ikaros and tumor suppression in acute lymphoblastic leukemia. Crit Rev Oncog. 2011; 16:3–12. [PubMed: 22150303]
35. Surveillance E, and End Results (SEER) Program. National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch; SEER*Stat Database: Incidence - SEER 13 Regs Research Data, Nov 2011 Sub, Vintage 2009 Pops (1992–2009) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969–2010 Counties. (www.seer.cancer.gov)released April 2012, based on the November 2011 submission [Accessed October 16, 2012]
36. Borkhardt A, Wuchter C, Viehmann S, Pils S, Teigler-Schlegel A, Stanulla M, et al. Infant acute lymphoblastic leukemia - combined cytogenetic, immunophenotypical and molecular analysis of 77 cases. Leukemia. 2002; 16:1685–90. [PubMed: 12200682]

37. Jansen MW, Corral L, van der Velden VH, Panzer-Grumayer R, Schrappe M, Schrauder A, et al. Immunobiological diversity in infant acute lymphoblastic leukemia is related to the occurrence and type of MLL gene rearrangement. Leukemia. 2007; 21:633–41. [PubMed: 17268512]

38. Olson JE, Shu XO, Ross JA, Pendergrass T, Robison LL. Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group. Am J Epidemiol. 1997; 145:58–67. [PubMed: 8982023]

Table 1

Selected characteristics of 770 childhood ALL cases.

	N (%)
Immunophenotype	
B cell	574 (74.6)
T cell	95 (12.3)
Unknown	101 (13.1)
Chromosomal numerical abnormalities	
None	96 (12.5)
Hyperdiploid (>50)	79 (10.3)
Other	122 (15.8)
Unknown	473 (61.4)
Diagnosis age (y)	
<1	32 (4.2)
1 – <2	61 (7.9)
2 – 5	422 (54.8)
6 – 10	174 (22.6)
11 – 14	81 (10.5)
Sex	
Male	424 (55.1)
Female	346 (44.9)
Race/ethnicity	
Non-Hispanic white	645 (83.8)
Non-Hispanic black	47 (6.1)
Hispanic	48 (6.2)
American Indian/Alaskan Native	7 (0.9)
Asian/Pacific Islander	20 (2.6)
Other	3 (0.4)
Birth weight (g)[1]	
≤3000	128 (16.6)
3001 – 3500	257 (33.4)
3501 – 4000	257 (33.4)
>4000	127 (16.5)
Maternal age at index child's birth (y)	
<20	69 (9.0)
20 – 24	185 (24.0)
25 – 29	285 (37.0)
30 – 34	163 (21.2)
≥35	68 (8.8)

ALL = acute lymphoblastic leukemia

[1] Birth weight not provided in maternal interview for one case

Table 2

Case-control comparisons for *ARID5B* and *IKZF1* single nucleotide polymorphisms in 770 childhood ALL cases and 384 healthy controls, for combined cases overall and for B lineage, B lineage hyperdiploid, and T lineage cases examined separately.

	Controls		Combined ALL Cases				B lineage ALL				B hyperdiploid ALL				T lineage ALL			
	N	**RAF**	**N**	**RAF**	**OR**	**95% CI**	**N**	**RAF**	**OR**	**95% CI**	**N**	**RAF**	**OR**	**95% CI**	**N**	**RAF**	**OR**	**95% CI**
***ARID5B* rs10821936**																		
TT	169		245		1.00		171		1.00		14		1.00		47		1.00	
TC	171	0.34	362	0.44	**1.46**	**1.12–1.91**	274	0.46	**1.58**	**1.19–2.11**	35	0.60	**2.47**	**1.28–4.76**	34	0.33	0.72	0.44–1.17
CC	43		160		**2.57**	**1.74–3.79**	127		**2.92**	**1.95–4.38**	30		**8.42**	**4.11–17.25**	14		1.17	0.59–2.32
Per C allele					**1.56**	**1.31–1.87**			**1.68**	**1.39–2.03**			**2.94**	**2.04–4.24**			0.96	0.69–1.34
P_{trend}[1]					**9×10^{-7}**				**7×10^{-8}**				**1×10^{-9}**				0.81	
***ARID5B* rs10994982**																		
GG	101		148		1.00		107		1.00		12		1.00		23		1.00	
GA	187	0.49	324	0.57	1.18	0.87–1.61	239	0.58	1.21	0.87–1.68	26	0.64	1.17	0.57–2.42	45	0.49	1.06	0.61–1.85
AA	95		243		**1.75**	**1.23–2.47**	187		**1.86**	**1.29–2.68**	32		**2.84**	**1.38–5.83**	21		0.97	0.50–1.87
Per A allele					**1.33**	**1.12–1.58**			**1.37**	**1.14–1.65**			**1.81**	**1.25–2.61**			0.99	0.71–1.36
P_{trend}[1]					**0.001**				**0.0007**				**0.001**				0.94	
***IKZF1* rs11978267**																		
AA	204		321		1.00		224		1.00		25		1.00		50		1.00	
AG	152	0.27	299	0.36	1.25	0.96–1.63	228	0.38	**1.37**	**1.03–1.81**	28	0.47	1.50	0.84–2.68	36	0.28	0.97	0.60–1.56
GG	28		110		**2.50**	**1.59–3.92**	90		**2.93**	**1.84–4.66**	21		**6.12**	**3.03–12.35**	8		1.17	0.50–2.71
Per G allele					**1.45**	**1.20–1.74**			**1.58**	**1.30–1.92**			**2.32**	**1.62–3.32**			1.03	0.72–1.47
P_{trend}[1]					**0.0001**				**5×10^{-6}**				**5×10^{-6}**				0.87	

ALL = acute lymphoblastic leukemia; CI = confidence interval; OR = odds ratio; RAF = Risk allele frequency

[1] *P*-value for trend calculated via Cochran-Armitage trend test

Table 3

Case-control comparisons for *ARID5B* and *IKZF1* single nucleotide polymorphisms in 770 childhood ALL cases and 384 healthy controls by sex.

	Controls		Combined ALL Cases				B lineage ALL				B hyperdiploid ALL				T lineage ALL			
	N	RAF	N	RAF	OR	95% CI	N	RAF	OR	95% CI	N	RAF	OR	95% CI	N	RAF	OR	95% CI
Females																		
***ARID5B* rs10821936**																		
TT	77		102		1.00		78		1.00		5		1.00		14		1.00	
TC	77	0.34	161	0.47	**1.58**	**1.06–2.36**	121	0.49	**1.55**	**1.01–2.37**	23	0.64	**4.60**	**1.66–12.72**	13	0.38	0.93	0.41–2.11
CC	22		83		**2.85**	**1.63–4.96**	70		**3.14**	**1.77–5.57**	18		**12.60**	**4.20–37.79**	6		1.50	0.52–4.36
Per C allele					**1.66**	**1.28–2.16**			**1.73**	**1.32–2.27**			**3.39**	**2.04–5.63**			1.16	0.68–1.96
P_{trend}[1]					**0.0001**				**7×10^{-5}**				**5×10^{-7}**				0.59	
***ARID5B* rs10994982**																		
GG	47		63		1.00		47		1.00		6		1.00		10		1.00	
GA	89	0.49	143	0.58	1.20	0.76–1.90	111	0.59	1.25	0.76–2.04	17	0.64	1.50	0.55–4.05	14	0.43	0.74	0.31–1.79
AA	41		111		**2.02**	**1.20–3.40**	91		**2.22**	**1.28–3.84**	17		**3.25**	**1.17–9.01**	6		0.69	0.23–2.06
Per A allele					**1.43**	**1.10–1.85**			**1.50**	**1.15–1.97**			**1.87**	**1.13–3.11**			0.82	0.47–1.42
P_{trend}[1]					**0.006**				**0.003**				**0.01**				0.48	
***IKZF1* rs11978267**																		
AA	92		149		1.00		110		1.00		17		1.00		19		1.00	
AG	74	0.27	128	0.35	1.07	0.73–1.57	102	0.37	1.15	0.77–1.73	13	0.47	0.95	0.43–2.08	12	0.24	0.79	0.36–1.72
GG	11		52		**2.92**	**1.45–5.88**	43		**3.27**	**1.60–6.70**	14		**6.89**	**2.68–17.70**	2		0.88	0.18–4.30
Per G allele					**1.42**	**1.08–1.87**			**1.53**	**1.14–2.04**			**2.24**	**1.39–3.62**			0.85	0.46–1.59
P_{trend}[1]					**0.01**				**0.004**				**0.0007**				0.62	
Males																		
***ARID5B* rs10821936**																		
TT	92		143		1.00		93		1.00		9		1.00		33		1.00	
TC	94	0.34	201	0.42	1.38	0.96–1.97	153	0.44	**1.61**	**1.09–2.37**	12	0.55	1.31	0.53–3.25	21	0.30	0.62	0.34–1.16
CC	21		77		**2.36**	**1.36–4.09**	57		**2.69**	**1.51–4.78**	12		**5.84**	**2.18–15.65**	8		1.06	0.43–2.63
Per C allele					**1.49**	**1.16–1.91**			**1.63**	**1.25–2.13**			**2.42**	**1.42–4.13**			0.87	0.57–1.34
P_{trend}[1]					**0.002**				**0.0003**				**0.0008**				0.53	

	Controls		Combined ALL Cases				B lineage ALL				B hyperdiploid ALL				T lineage ALL			
	N	**RAF**	**N**	**RAF**	**OR**	**95% CI**	**N**	**RAF**	**OR**	**95% CI**	**N**	**RAF**	**OR**	**95% CI**	**N**	**RAF**	**OR**	**95% CI**
$P_{\text{Female vs. Male}}$[2]					0.54				0.81				0.37				0.42	
***ARID5B* rs10994982**																		
GG	54		85		1.00		60		1.00		6		1.00		13		1.00	
GA	98	0.49	181	0.56	1.17	0.77–1.79	128	0.56	1.18	0.75–1.85	9	0.65	0.83	0.28–2.45	31	0.52	1.31	0.64–2.72
AA	54		132		1.55	0.98–2.47	96		1.60	0.97–2.63	15		2.50	0.90–6.92	15		1.15	0.50–2.65
Per A allele					1.25	0.99–1.57			1.27	0.99–1.63			**1.77**	**1.03–3.07**			1.07	0.71–1.60
P_{trend}[1]					0.06				0.06				**0.04**				0.75	
$P_{\text{Female vs. Male}}$[2]					0.45				0.49				0.89				0.45	
***IKZF1* rs11978267**																		
AA	112		172		1.00		114		1.00		8		1.00		31		1.00	
AG	78	0.27	171	0.36	1.43	1.00–2.04	126	0.38	**1.59**	**1.08–2.33**	15	0.48	**2.69**	**1.09–6.66**	24	0.30	1.11	0.61–2.04
GG	17		58		**2.22**	**1.23–4.01**	47		**2.72**	**1.47–5.01**	7		**5.77**	**1.85–17.94**	6		1.28	0.46–3.51
Per G allele					**1.47**	**1.14–1.89**			**1.63**	**1.24–2.13**			**2.43**	**1.40–4.22**			1.12	0.73–1.74
P_{trend}[1]					**0.003**				**0.0004**				**0.001**				0.60	
$P_{\text{Female vs. Male}}$[2]					0.87				0.95				0.83				0.48	

ALL = acute lymphoblastic leukemia; CI = confidence interval; OR = odds ratio; RAF = Risk allele frequency

[1] *P*-value for trend calculated via Cochran-Armitage trend test

[2] *P*-value for 1 degree of freedom test of interaction from the allelic model

Table 4

Genotype frequencies and case-only allelic interaction odds ratios for *ARID5B* and *IKZF1* single nucleotide polymorphisms and birth weight and maternal age categories, respectively, in 770 childhood ALL cases.

	***ARID5B* rs10821936**						
	TT	**CT**	**CC**	**RAF**	**IOR**	**95% CI**	***P*[1]**
Birth weight (g)[2]							
≤3000	45	45	25	0.41	0.92	0.60–1.42	
3001 – 3500	83	119	51	0.44	1.00		0.82
3501 – 4000	76	130	49	0.45	1.06	0.77–1.47	
>4000	37	60	29	0.47	1.16	0.77–1.73	
Maternal age at index child's birth (y)							
<20	19	32	17	0.49	1.13	0.69–1.86	
20 – 24	74	77	34	0.39	**0.68**	**0.48–0.97**	
25 – 29	82	141	60	0.46	1.00		0.08
30 – 34	46	78	39	0.48	1.09	0.76–1.57	
≥35	24	34	10	0.40	0.72	0.44–1.19	
	***ARID5B* rs10994982**						
	GG	**GA**	**AA**	**RAF**	**IOR**	**95% CI**	***P*[1]**
Birth weight (g)[2]							
≤3000	22	45	38	0.58	1.21	0.77–1.91	
3001 – 3500	48	111	76	0.56	1.00		0.51
3501 – 4000	49	105	86	0.58	1.10	0.78–1.53	
>4000	27	58	35	0.53	0.84	0.56–1.27	
Maternal age at index child's birth (y)							
<20	10	31	25	0.61			
20 – 24	51	60	52	0.50			
25 – 29	46	132	89	0.58	*3*		
30 – 34	30	65	58	0.59			
≥35	11	36	19	0.56			

	IKZF1 rs11978267						
	AA	**AG**	**GG**	**RAF**	**IOR**	**95% CI**	***P*[1]**
Birth weight (g)[2]							
≤3000	35	54	18	0.42	1.50	0.96–2.35	
3001 – 3500	109	94	37	0.35	1.00		0.12
3501 – 4000	118	93	33	0.33	0.88	0.63–1.23	
>4000	51	53	19	0.37	1.12	0.74–1.70	
Maternal age at index child's birth (y)							
<20	26	30	10	0.38	1.17	0.70–1.93	
20 – 24	85	65	20	0.31	0.79	0.55–1.14	
25 – 29	121	107	42	0.35	1.00		0.15
30 – 34	67	65	25	0.37	1.08	0.74–1.56	
≥35	22	32	13	0.43	1.53	0.93–2.53	

ALL = acute lymphoblastic leukemia; CI = confidence interval; IOR = interaction odds ratio; RAF=risk allele frequency

[1] *P*-values for interactions (with 3 (birth weight) and 4 (maternal age) degrees of freedom, respectively)

[2] Birth weight models adjusted for gestational age. Models excluded 17 cases that were born in a set of multiples (e.g., twins).

[3] Allelic IORs not reported because the proportional hazards assumption was violated (P=0.02). Genotypic IORs are: <20 y: GA: 1.08 (0.49–2.38), AA: 1.29 (0.57–2.92); 20–24 y: GA: 0.41 (0.25–0.68), AA: 0.53 (0.31–0.89); 25–29 y: reference; 30–34 y: GA: 0.76 (0.44–1.31), AA: 1.00 (0.57–1.76); ≥35 years: GA: 1.14 (0.54–2.42), AA: 0.89 (0.39–2.03); $P_{interaction}$=0.02.