Received: 27 February 2024 | Revised: 6 April 2024 | Accepted: 22 April 2024

DOI: 10.1002/gcc.23242

RESEARCH ARTICLE

WILEY

Constitutional and acquired genetic variants in *ARID5B* in pediatric B-cell precursor acute lymphoblastic leukemia

Charlotte Ragnarsson[1,2] | Minjun Yang[1] | Larissa Helena Moura-Castro[1] | Efe Aydın[1] | Rebeqa Gunnarsson[1] | Linda Olsson-Arvidsson[1,3] | Henrik Lilljebjörn[1] | Thoas Fioretos[1,3] | Nicolas Duployez[4] | Marketa Zaliova[5,6] | Jan Zuna[5,6] | Anders Castor[2] | Bertil Johansson[1,3] | Kajsa Paulsson[1]

[1]Division of Clinical Genetics, Department of Laboratory Medicine, Lund University, Lund, Sweden

[2]Department of Paediatrics, Skåne University Hospital, Lund University, Lund, Sweden

[3]Department of Clinical Genetics, Pathology, and Molecular Diagnostics, Office for Medical Services, Region Skåne, Lund, Sweden

[4]Laboratory of Haematology, Centre Hospitalier Universitaire (CHU) Lille, University of Lille, INSERM Unité 1277 Canther, Lille, France

[5]Department of Paediatric Haematology and Oncology, Second Faculty of Medicine, Charles University/University Hospital Motol, Prague, Czech Republic

[6]Childhood Leukaemia Investigation Prague (CLIP), Prague, Czech Republic

Correspondence
Kajsa Paulsson, Division of Clinical Genetics, BMC C13, SE-221 85 Lund, Sweden.
Email: kajsa.paulsson@med.lu.se

Funding information
The Swedish Research Council; Cancerfonden; Barncancerfonden; National Institute for Cancer Research funded by the European Union Next Generation EU; Governmental Funding of Clinical Research within the National Health Service

Abstract

Constitutional polymorphisms in *ARID5B* are associated with an increased risk of developing high hyperdiploid (HeH; 51–67 chromosomes) pediatric B-cell precursor acute lymphoblastic leukemia (BCP ALL). Here, we investigated constitutional and somatic *ARID5B* variants in 1335 BCP ALL cases from five different cohorts, with a particular focus on HeH cases. In 353 HeH ALL that were heterozygous for risk alleles and trisomic for chromosome 10, where *ARID5B* is located, a significantly higher proportion of risk allele duplication was seen for the SNPs rs7090445 ($p = 0.009$), rs7089424 ($p = 0.005$), rs7073837 ($p = 0.03$), and rs10740055 ($p = 0.04$). Somatic *ARID5B* deletions were seen in 16/1335 cases (1.2%), being more common in HeH than in other genetic subtypes (2.2% vs. 0.4%; $p = 0.002$). The expression of *ARID5B* in HeH cases with genomic deletions was reduced, consistent with a functional role in leukemogenesis. Whole-genome sequencing and RNA-sequencing in HeH revealed additional somatic events involving *ARID5B*, resulting in a total frequency of 3.6% of HeH cases displaying a somatic *ARID5B* aberration. Overall, our results show that both constitutional and somatic events in *ARID5B* are involved in the leukemogenesis of pediatric BCP ALL, particularly in the HeH subtype.

KEYWORDS
ARID5B variants, B-cell precursor acute lymphoblastic leukemia, constitutional, high hyperdiploidy, pediatric, somatic

1 | INTRODUCTION

Constitutional single nucleotide polymorphisms (SNPs) in the *ARID5B* gene at 10q21.2, which codes for a member of the AT-rich interactive domain (ARID) protein family involved in chromatin-remodeling,[1,2] are associated with an increased risk of developing childhood B cell precursor acute lymphoblastic leukemia (BCP ALL), particularly of the high hyperdiploid (HeH; 51–67 chromosomes) subtype[3–11] as well as

with an increased risk of relapse of pediatric ALL.[6,7,9] Two recent publications have reported that *ARID5B* expression plays an important role in regulating hematopoiesis at the pre-B-stage.[12,13] In a mouse model, Chalise et al[12] showed that deletion of *ARID5B* led to a significant reduction in immature B cells. Furthermore, a decrease in *ARID5B* expression was observed in leukemic blast cells from BCP ALL patients in comparison with B-cells from nonleukemic individuals. Also, lower *ARID5B* expression was associated with decreased survival in BCP ALL patients.[12] Goodings et al.[13] found that transgenic mice with overexpression of *Arid5b* had a reduction in circulating B-cells, immature and mature B-cells, mainly driven by a lower lymphocyte count, and a decrease of follicular B cells in the spleen.

The HeH subtype accounts for approximately 25% of BCP ALL cases. It is characterized by a nonrandom gain of chromosomes and is associated with a favorable prognosis.[14] Chromosome 10, containing the *ARID5B* gene, is trisomic in around 70% of the cases. We have previously reported that HeH cases that are heterozygous for the rs7089424 risk allele more frequently have a gain of the chromosome 10 homolog that carries this specific allele.[5,7] Furthermore, we have reported two cases with somatic deletions of *ARID5B* among 74 pediatric HeH ALL cases.[5] Here, we further address the role of constitutional and somatic *ARID5B* variants by investigating a large series of pediatric ALL cases, including both HeH and non-HeH cases, for an association between the risk alleles of *ARID5B* and trisomy 10. We also determine the frequency of acquired somatic variants in *ARID5B*.

2 | MATERIALS AND METHODS

2.1 | Patients

A total of 1335 BCP ALL cases from five different cohorts were included in the investigation, of which 590 were HeH. Cohort 1 comprises 268 HeH ($n = 113$) and non-HeH ($n = 155$) cases and is an expansion from a previously published cohort.[15–17] Cohort 2, The Therapeutically Applicable Research to Generate Effective Treatments (TARGET, dbGAP accession number phs000464) comprises 706 cases (HeH; $n = 116$, non-HeH; $n = 590$) genotyped using either the Affymetrix Genome-Wide Human SNP Array 6.0 or whole genome sequencing (WGS) based on the Complete Genomics technology. Cohort 3 was obtained from St. Jude Cloud[18–21] and comprised 141 HeH cases genotyped using Illumina (WGS). Cohort 4[22] comprises 124 HeH cases that were genotyped by Affymetrix CytoScan HD array. Cohort 5[23] comprises 96 HeH cases genotyped by Illumina BeadChips. Cases were assigned to genetic subtypes based on having a primary fusion gene or chromosomal copy number changes indicative of HeH, near-haploidy, or low hypodiploidy. The B-other group comprised BCP ALLs without HeH, near-haploidy, low hypodiploidy, intrachromosomal amplification of chromosome 21, *BCR*::*ABL1*, *ETV6*::*RUNX1*, and *TCF3*::*PBX1* fusions, and *KMT2A* rearrangements. Informed consent was obtained according to the Declaration of Helsinki, and the study was approved by the Swedish Ethical Review Authority.

2.2 | SNP array data analysis

Log R ratio (LRR) and B allele frequency (BAF) of SNP array data from Illumina (.idat files) and Affymetrix (.CEL files) intensity files were analyzed by GenomeStudio (v2.0, Illumina, San Diego, CA) and Affymetrix Analysis Power Tools (v2.10.0, Thermo Fisher Scientific Inc., Waltham, MA), respectively. The copy-number of chromosome 10 was called with Tumor Aberration Prediction Suite (TAPS)[24] and manually reviewed in GenomeStudio or Chromosome Analysis Suite. The genotype and relative allele frequencies of risk SNPs in *ARID5B* were ascertained by the B allele frequency (BAF). Only risk SNPs were included, and no imputation was performed.

2.3 | WGS data analysis

WGS data were available from 217 cases (38 cases from cohort 1, 38 cases from cohort 2 and 141 cases from cohort 3) of which all were HeH. For Complete Genomics data generated by TARGET, the detection of constitutional and somatic variants of *ARID5B* was done with the Complete Genomics Cancer Sequencing v2.0 pipeline (Complete Genomics). Somatic variants were further filtered for somatic Score ≥0 and the number of unique reads for the mutated allele >10. For WGS data sequenced with the Illumina platform, sequencing reads were aligned to the UCSC human reference genome (hg19) using the Burrows–Wheeler Aligner (BWA, v0.7.17). Duplicate reads marking and local realignment were performed by the Genome Analysis Toolkit (GATK).[25] Constitutional variants of matched tumor/normal pairs were called by GATK HaplotypeCaller. Somatic variants were identified by the GDC DNA-Seq analysis pipeline (https://docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/DNA_Seq_Variant_Calling_Pipeline/), and structural variants of *ARID5B* were called by Manta and Gridss2.[26,27] To investigate the copy number of chromosome number 10, read counts of constitutional mutation sites were normalized to the sequencing depth; the tumor sample's Log R ratio (LRR) was then calculated by the log-odds ratio of the variant allele count in the tumor versus in the normal. BAF was defined by the reference allele count versus the total sequencing depth of the constitutional variant site in the tumor sample. Copy number alterations were called using TAPS.[24] Only calls that passed the somatic filtering process for each caller were retained. Furthermore, somatic variants were validated if they were identified by both callers, and their breakpoints were supported by at least three reads from each caller.

2.4 | RNA sequencing

RNA sequencing (RNA-seq) data from cohort 1 have been published ($n = 52$).[28] For RNA-seq data from cohort 2 ($n = 25$), methods are available at the TARGET project portal TARGET (dbGAP accession number phs000464). RNA-seq reads were mapped to the human reference genome (hg19) using the STAR (version 2.6.1d) 2-pass mapping pipeline. All included cases were HeH. Gene expression levels

were estimated using the GDC mRNA quantification analysis pipeline (https://docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/Expression_mRNA_Pipeline/). Read counts of coding genes were converted into TPM (transcripts per million) value. To control for ploidy effects, cases with disomy and trisomy 10 were considered separately.

2.5 | Fluorescence in situ hybridization (FISH)

FISH analysis was performed on 22 metaphase cells according to standard methods using whole chromosome paint probes (Applied Spectral Imaging, Carlsbad, CA) for chromosomes X (Cy3 red) and 10 (Aqua blue) to validate the t(X;10)(q23;q21.2) in one case with *STAG2*::*ARID5B* fusion detected by RNA-seq.

2.6 | Statistical analysis

For the evaluation of constitutional variants in *ARID5B*, a binomial one-sided test was performed to analyze whether there was a higher frequency of duplicated risk alleles in heterozygous cases trisomic for chromosome 10. To compare the proportions of deletions in the HeH cohort and the other genetic subtypes, Fisher's exact two-sided test was used. Fisher's exact two-sided test was also used to evaluate if there was a significant difference in gender distribution in HeH cases with deletions in the *ARID5B* region compared to HeH without such deletions. Mann Whitney *U* test was used to ascertain if there was a significant difference in age distribution and *ARID5B* gene expression of HeH cases with deletions in the *ARID5B* region compared to those without such deletions. *p*-values <0.05 were considered significant.

3 | RESULTS

3.1 | Preferred risk allele duplication

The distribution of the constitutional risk alleles rs7090445, rs7089424, rs7073837, and rs10740055 of *ARID5B* were investigated in a total of 353 HeH ALL heterozygous for the risk alleles and trisomic for chromosome 10. Data from cohort 1 have been previously published and were therefore not included in this analysis.[5] The three different cohorts (116 cases from cohort 2, 141 from cohort 3, and 96 from cohort 5) included a total of 92 cases informative for rs7090445, 92 cases for rs7089424, 119 cases for rs7073837, and 66 cases for rs10740055. All four risk SNPs showed a significantly higher proportion of risk allele duplication (one-sided binominal test; $p < 0.05$) (Table 1).

3.2 | Somatic variants involving *ARID5B* in HeH and non-HeH pediatric BCP ALL

A total of 16 deletions (1.2%) overlapping *ARID5B* were found among the 1335 HeH and non-HeH BCP ALL (Table 2). The deletions covered different parts of *ARID5B*, with no minimally overlapping region (Figure 1). All deletions except for one (non-HeH) were hemizygous, with sizes ranging from 7.9 to 257 kb (median size 73.2 kb; Table 2, Figure 1). Of the 16 cases with 10q21.2 deletions, paired diagnostic and relapse samples were available in two cases (cohort 2; cases designated PARBRK and PARTKL), which both harbored the deletions at both time points. Somatic *ARID5B* deletions were more common in HeH cases (13/590; 2.2%) than in cases from other genetic subtypes (3/745; 0.4%) ($p = 0.002$ Fisher's exact two-sided test). No significant difference in gender was seen in informative deleted vs. nondeleted cases: 2/116 (1.7%) of males and 4/99 (4.0%) of females had *ARID5B* deletions ($p = 0.4$, Fisher's exact two-sided test). Furthermore, no significant difference in age was seen in 221 informative cases (Mann–Whitney *U* test, $p = 0.5$). Seven of the 13 HeH cases harboring a deletion within the *ARID5B* region had disomy 10 and 6 trisomy 10. Regardless of chromosomal copy number, there was always one copy deleted in these cases (Table 2). For the three cases from other genetic subtypes, which all had disomy 10, 1 had a homozygous deletion – constituting two separate deletions – and 2 had hemizygous deletions (Table 2).

WGS data were available for 217 HeH cases, allowing the identification of other types of somatic variants. This analysis revealed one translocation t(10;13)(q21.2;12.2) in cohort 1 (*ARID5B*::*PAN3*

TABLE 1 Risk allele duplication in high hyperdiploid pediatric B-cell acute lymphoblastic leukemia cases with trisomy 10.

Allele	No of cases with risk allele duplication (%)	No of cases with non-risk allele duplication (%)	Total no of cases	*p*-value*
rs7090445 (C = risk allele)	59 (64)	33 (36)	92	0.009
rs7089424 (C = risk allele)	60 (65)	32 (35)	92	0.005
rs7073837 (A = risk allele)	72 (61)	47 (39)	119	0.03
rs10740055 (C = risk allele)	42 (64)	24 (36)	66	0.04

*One-sided binominal test.

TABLE 2 Somatic deletions involving ARID5B in 16 cases of pediatric B-cell precursor acute lymphoblastic leukemias.

Case	Deleted region	Size (bp)	Gender	Subtype	No of copies chromosome 10	No of copies left in deleted region
Cohort 1 L1	chr10: 63516353–63772939	256 586	M	HeH	2	1
Cohort 1 L2	chr10: 63799904–63843012	43 108	M	HeH	2	1
Cohort 2 PANMEN	chr10: 63632846–63740912	108 066	F	HeH	3	2
Cohort 2 PARBRK	chr10: 63832278–63862428	30 150	F	HeH	3	2
Cohort 2 PARTKL	chr10: 63722731–63783691	60 960	F	HeH	2	1
Cohort 2 PASFTL	chr10: 63594799–63 774 235	179 436	F	HeH	2	1
Cohort 3 SJBALL030103	chr10: 63808800–63816708	7908	M	HeH	3	2
Cohort3 SJBALL030644	chr10: 63784841–63833104	48 263	F	HeH	3	2
Cohort 3 SJHYPER006	chr10: 63661111–63747960	86 849	M	HeH	3	2
Cohort 3 SJHYPER043	chr10: 63684774–63779925	95 151	M	HeH	3	2
Cohort 3 SJHYPER108	chr10: 63657737–63870072	212 335	M	HeH	2	1
Cohort 4 D39	chr10: 63710103–63736912	26 809	F	HeH	2	1
Cohort 4 D72	chr10: 63666523–63751978	85 455	M	HeH	2	1
Cohort 2 PANWNG	6370224–63706250	36 026	F	B-other	2	1
Cohort 2 PAPCSR	63684083–63814817	130 734	F	*ETV6::RUNX1*	2	0-1
Cohort 2 PAPFIB	6366690–63714778	48 088	M	B-other	2	1

Abbreviations: bp, base pair; F, female; HeH, high hyperdiploid; M, male; N/A, data not available (no relapse or no sample available); NK, not known.

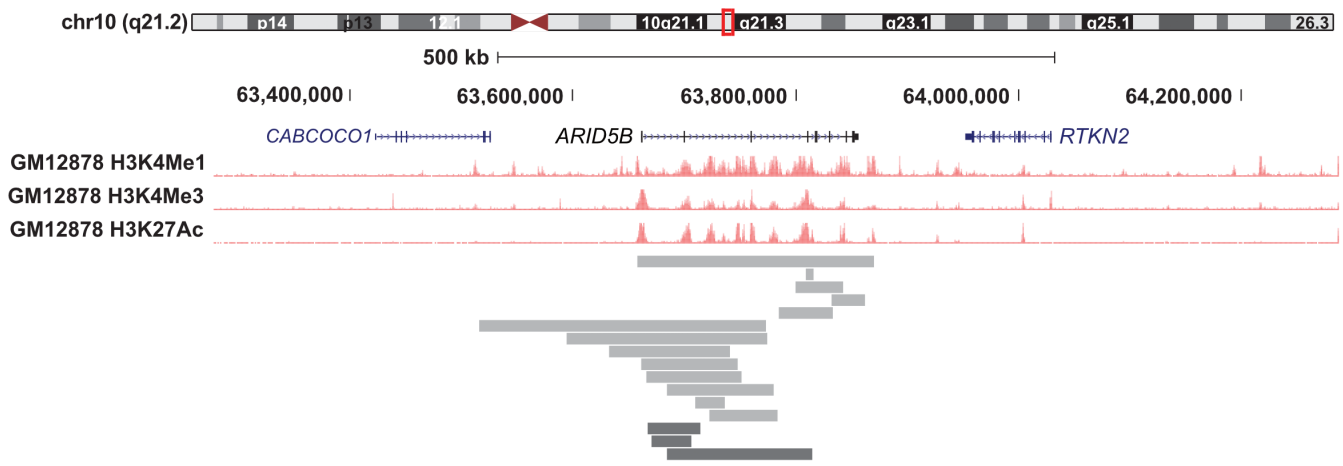

FIGURE 1 Sixteen somatic deletions in 10q21.1 involving *ARID5B* in pediatric B-cell precursor acute lymphoblastic leukemia (ALL). The top panel shows the genomic location of the *ARID5B* gene locus on chromosome 10, followed by the histone modifications ChIP-seq data, including H3K4me1, H3K4me3 and H3K27ac based on the GM12878 cell line data from UCSC. Deletions are shown as gray (high hyperdiploid ALL) or black (non-high hyperdiploid ALL) squares. Each square represents a deletion in a separate case. No minimally deleted region was detected.

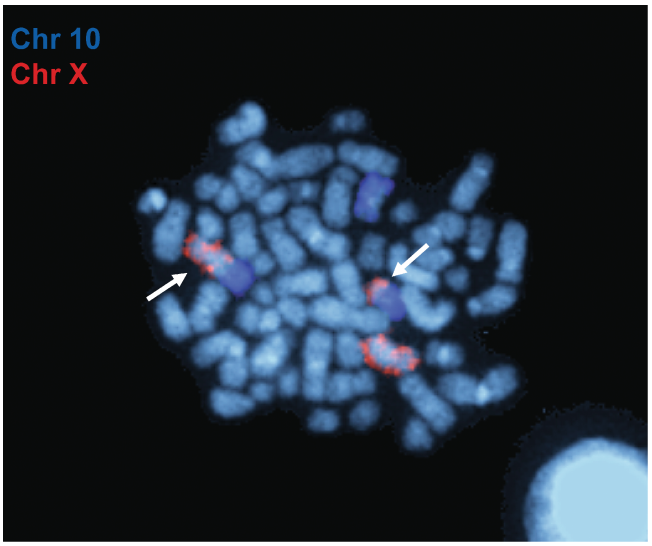

FIGURE 2 Metaphase fluorescence in situ hybridization of the t(X;10)(q25;q21.2) involving the *STAG2* and *ARID5B* genes. Chromosome 10 is labeled blue (Aqua) and chromosome X red (Cy3). The derivative chromosomes are marked with arrows.

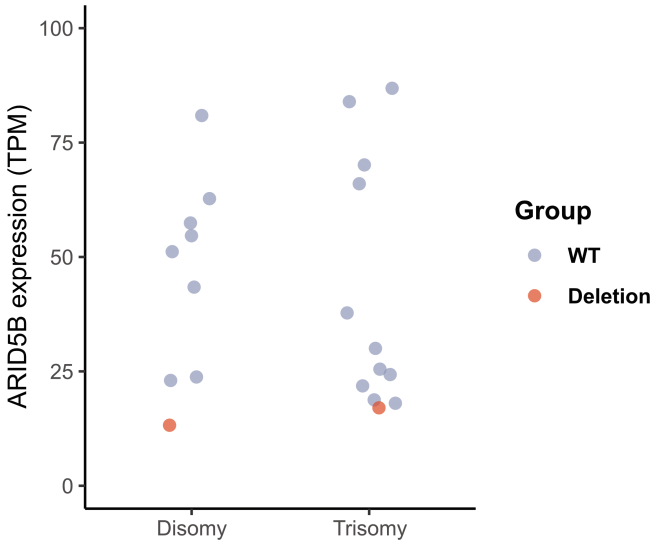

FIGURE 3 RNA expression data were available from two cases with *ARID5B* deletion: one disomic and one trisomic for chromosome 10. Both cases showed the lowest expression of this gene compared to the other cases with the same chromosome 10 copy number (among a total of 8 cases disomic and 11 cases trisomic for chromosome 10), indicating that cases deletions targeting *ARID5B* had reduced expression of the *ARID5B* gene. WT, wildtype (no deletion).

chr10:63721264, chr13:28728190), one missense mutation in cohort 2 (chr10:63817051) and one splice site insertion in cohort 3 (chr10:63661916-63 661 917) giving a frequency of 3/217 cases (1.4%). The proportion of reads covering the *ARID5B::PAN3* breakpoint was 23.8% based on the structural variation caller Manta, which is in line with it involving one of three copies of chromosome 10. However, since chromosome 13 in this case was a uniparental disomy (UPD) and loss of heterozygosity was seen across the whole chromosome, the translocation must have arisen after the UPD. In addition, one more translocation involving *ARID5B* was identified in cohort 1 by RNA sequencing (*STAG2::ARID5B*; fusion of exon 1 of *STAG2* and exon 8 of *ARID5B*). A corresponding t(X;10)(q25;q21.2) was seen by FISH in 20/22 metaphases (Figure 2), with the remaining two metaphases showing only trisomy X, indicating that the translocation was subclonal. None of the two translocations was predicted to result in a functional fusion gene.

3.3 | RNA expression data from HeH cases

RNA expression data were available from two cases with *ARID5B* deletions: one disomic and one trisomic for chromosome 10. These two cases had the lowest expression of this gene of all cases with the same chromosome 10 copy number (among a total of 8 disomic and 11 trisomic cases) (Figure 3).

4 | DISCUSSION

Here, we address the role of constitutional and somatic *ARID5B* variants in pediatric BCP ALL by investigating a large series of such cases, including both HeH and non-HeH. Certain alleles of *ARID5B* show a very robust association with increased risk of pediatric HeH BCP ALL over different populations. Previous studies have shown that the risk allele lowers *ARID5B* expression[7] and that dysregulation of this gene affects B-cell hematopoiesis.[12,13] Why this results in increased risk specifically for HeH ALL is not known. A characteristic feature of the HeH subtype is the nonrandom gain of chromosomes, with an extra copy of chromosome 10, that carries *ARID5B*, being seen in ~70% of cases.[5,17] We have previously reported that HeH ALLs that are constitutionally heterozygous for *ARID5B* risk alleles and with an acquired trisomy 10 more commonly gain the chromosome 10 homolog carrying the risk allele.[7] Here, we validate this finding in a much larger cohort. This indicates opposing evolutionary pressures in the clonal evolution of HeH ALL, where the gain of chromosome 10 overall provides a selective advantage, but raising the expression of *ARID5B* gives a competitive disadvantage, resulting in preferential duplication of the homolog carrying the risk allele.

We further studied somatic variants involving *ARID5B* and found recurrent somatic deletions targeting *ARID5B* in pediatric BCP ALL, at an overall frequency of 1.2%. These were significantly more frequent in HeH than in non-HeH cases, again underscoring the specific role of *ARID5B* in the leukemogenesis of HeH ALL. Notably, although all deletions targeted coding regions of *ARID5B*, they covered different parts of the gene, with no minimally deleted region, suggesting that the functional outcome would be downregulation of *ARID5B* expression. None of the 6 cases with trisomy 10 and *ARID5B* deletions had duplication of the deleted chromosome. This could either be because the deletions are late events occurring after the gain of trisomy 10, or because the other chromosomal homologue was duplicated. Thus,

larger cohorts are needed to determine whether somatic *ARID5B* deletions are early or late events. However, the two translocations involving *ARID5B* could be demonstrated to occur after UPD13 and trisomy X, respectively, indicating that they occurred after the hyperdiploidy. We identified two cases with *ARID5B* deletions and available RNA expression data; both of these displayed the lowest expression of this gene within their respective chromosome 10 copy number group. Interestingly, all except one of the deletions in 10q21.2 were hemizygous, indicating that lowering *ARID5B* expression, rather than removing it completely, is beneficial for the leukemic blast cells.

Other somatic variants were also found in a subset of HeH BCP ALL cases analyzed by WGS, comprising two translocations, one missense mutation, and one splice site insertion. The functional outcome of ARID5B for both translocations appears to be a loss of expression, in line with low levels of *ARID5B* providing a selective advantage. Together with the somatic deletions, the overall frequency of somatic events targeting *ARID5B* amounts to approximately 3.6%.

Overall, our results are consistent with an interplay between constitutional *ARID5B* risk alleles and acquired trisomy 10 in HeH ALL. Furthermore, we show that somatic events involving *ARID5B*, predominately deletions, are recurrent in pediatric BCP ALL and enriched in HeH cases. Constitutional *ARID5B* variants have been associated with an elevated risk of relapse as well as development of resistance to treatment.[9] Therefore, future studies should address whether somatic variants in *ARID5B* could be useful as markers for treatment stratification in BCP ALL.

AUTHOR CONTRIBUTIONS

CR and KP designed the study. CR, MY, LHM-C, EA, RG, LO-A, HL, TF, ND, MZ, JZ, AC, BJ, and KP analyzed data. KP supervised the study. CR, MY, and KP wrote the article with input from all authors.

ACKNOWLEDGMENTS

The results published here are in part based on data from the Therapeutically Applicable Research to Generate Effective Treatments (http://ocg.cancer.gov/programs/target). WGS data for pediatric tumor samples used for analysis in this study were obtained from St. Jude Cloud[18–21] (https://www.stjude.cloud).

FUNDING INFORMATION

This study was supported by grants from the Swedish Childhood Cancer Fund, grant numbers ST2021-0003 (AC), PR2020-0033 (MY), TJ2020-0024 (MY), PR2021-0005 (BJ) and PR2021-0016 (KP); the Swedish Cancer Fund, grant numbers 23 2694 Pj (BJ) and 20 0792 PjF and 22 2062 Pj (KP); Governmental Funding of Clinical Research within the National Health Service (BJ, KP); National Institute for Cancer Research, grant number LX22NPO5102, funded by the European Union–Next Generation EU (MZ and JZ); and the Swedish Research Council, grant numbers 2020-01164 (BJ) and 2020-00997 (KP).

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interest.

DATA AVAILABILITY STATEMENT

The Whole Genome Sequencing (WGS) and RNA Sequencing (RNA-seq) datasets from Lund University are deposited in the European Genome-phenome Archive (EGA), under the accession numbers EGAS00001007052 (WGS data) and EGAS00001001795 (RNA-seq data), respectively. WGS data pertaining to pediatric tumor samples analyzed in this study were sourced from the St. Jude Cloud, accessible via https://www.stjude.cloud/. Additionally, SNP array, WGS, and RNA-seq datasets produced by the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) initiative can be found at the Genomic Data Commons (GDC) portal, https://portal.gdc.cancer.gov/projects, under the accession code phs000464. The remaining SNP array datasets utilized in this study are hosted by SciLifeLab, available at https://doi.org/10.17044/scilifelab.21989795.v1.

ORCID

Charlotte Ragnarsson https://orcid.org/0009-0009-7244-9521
Larissa Helena Moura-Castro https://orcid.org/0000-0001-9063-5592
Henrik Lilljebjörn https://orcid.org/0000-0001-8703-1173
Marketa Zaliova https://orcid.org/0000-0002-1639-7124
Kajsa Paulsson https://orcid.org/0000-0001-7950-222X

REFERENCES

1. Wang X, Nagl NG, Wilsker D, et al. Two related ARID family proteins are alternative subunits of human SWI/SNF complexes. *Biochem J*. 2004;383(Pt 2):319-325.
2. Lin C, Song W, Bi X, et al. Recent advances in the ARID family: focusing on roles in human cancer. *Onco Targets Ther*. 2014;7:315-324.
3. Trevino LR, Yang W, French D, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nat Genet*. 2009;41(9):1001-1005.
4. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet*. 2009;41(9):1006-1010.
5. Paulsson K, Forestier E, Lilljebjörn H, et al. Genetic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*. 2010;107(50):21719-21724.
6. Xu H, Yang W, Perez-Andreu V, et al. Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. *J Natl Cancer Inst*. 2013;105(10):733-742.
7. Studd JB, Vijayakrishnan J, Yang M, Migliorini G, Paulsson K, Houlston RS. Genetic and regulatory mechanism of susceptibility to high-hyperdiploid acute lymphoblastic leukaemia at 10p21.2. *Nat Commun*. 2017;8:14616.
8. Xu H, Cheng C, Devidas M, et al. ARID5B genetic polymorphisms contribute to racial disparities in the incidence and treatment outcome of childhood acute lymphoblastic leukemia. *J Clin Oncol*. 2012;30(7):751-757.
9. Xu H, Zhao X, Bhojwani D, et al. ARID5B influences antimetabolite drug sensitivity and prognosis of acute lymphoblastic leukemia. *Clin Cancer Res*. 2020;26(1):256-264.
10. Zhao X, Qian M, Goodings C, et al. Molecular mechanisms of ARID5B-mediated genetic susceptibility to acute lymphoblastic leukemia. *J Natl Cancer Inst*. 2022;114(9):1287-1295.
11. Chokkalingam AP, Hsu LI, Metayer C, et al. Genetic variants in ARID5B and CEBPE are childhood ALL susceptibility loci in Hispanics. *Cancer Causes Control*. 2013;24(10):1789-1795.

12. Chalise JP, Ehsani A, Lemecha M, et al. ARID5B regulates fatty acid metabolism and proliferation at the pre-B cell stage during B cell development. *Front Immunol*. 2023;14:1170475.
13. Goodings C, Zhao X, McKinney-Freeman S, Zhang H, Yang JJ. ARID5B influences B-cell development and function in mouse. *Haematologica*. 2023;108(2):502-512.
14. Paulsson K, Johansson B. High hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosomes Cancer*. 2009;48(8):637-660.
15. Paulsson K, Lilljebjörn H, Biloglav A, et al. The genomic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Nat Genet*. 2015;47(6):672-676.
16. Olsson L, Castor A, Behrendtz M, et al. Deletions of IKZF1 and SPRED1 are associated with poor prognosis in a population-based series of pediatric B-cell precursor acute lymphoblastic leukemia diagnosed between 1992 and 2011. *Leukemia*. 2014;28(2):302-310.
17. Woodward EL, Yang M, Moura-Castro LH, et al. Clonal origin and development of high hyperdiploidy in childhood acute lymphoblastic leukaemia. *Nat Commun*. 2023;14(1):1658.
18. McLeod C, Gout AM, Zhou X, et al. St. Jude cloud: a pediatric cancer genomic data-sharing ecosystem. *Cancer Discov*. 2021;11(5):1082-1099.
19. Newman S, Nakitandwe J, Kesserwan CA, et al. Genomes for kids: the scope of pathogenic mutations in pediatric cancer revealed by comprehensive DNA and RNA sequencing. *Cancer Discov*. 2021;11(12):3008-3027.
20. Rusch M, Nakitandwe J, Shurtleff S, et al. Clinical cancer genomic profiling by three-platform sequencing of whole genome, whole exome and transcriptome. *Nat Commun*. 2018;9(1):3962.
21. Downing JR, Wilson RK, Zhang J, et al. The pediatric cancer genome project. *Nat Genet*. 2012;44(6):619-622.
22. Duployez N, Abou Chahla W, Lejeune S, et al. Detection of a new heterozygous germline ETV6 mutation in a case with hyperdiploid acute lymphoblastic leukemia. *Eur J Haematol*. 2018;100(1):104-107.
23. Zaliova M, Hovorkova L, Vaskova M, Hrusak O, Stary J, Zuna J. Slower early response to treatment and distinct expression profile of childhood high hyperdiploid acute lymphoblastic leukaemia with DNA index < 1.16. *Genes Chromosomes Cancer*. 2016;55(9):727-737.
24. Rasmussen M, Sundstrom M, Goransson Kultima H, et al. Allele-specific copy number analysis of tumor samples with aneuploidy and tumor heterogeneity. *Genome Biol*. 2011;12(10):R108.
25. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.
26. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*. 2016;32(8):1220-1222.
27. Cameron DL, Baber J, Shale C, et al. GRIDSS2: comprehensive characterisation of somatic structural variation using single breakend variants and structural variant phasing. *Genome Biol*. 2021;22(1):202.
28. Lilljebjörn H, Henningsson R, Hyrenius-Wittsten A, et al. Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. *Nat Commun*. 2016;7:11790.

How to cite this article: Ragnarsson C, Yang M, Moura-Castro LH, et al. Constitutional and acquired genetic variants in *ARID5B* in pediatric B-cell precursor acute lymphoblastic leukemia. *Genes Chromosomes Cancer*. 2024;63(5):e23242. doi:10.1002/gcc.23242