A germ-line deletion of APOBEC3B does not contribute to subtype-specific childhood acute lymphoblastic leukemia etiology

Approximately 4/100,000 children are diagnosed with acute lymphoblastic leukemia (ALL) in the United States annually. Early life exposures related to immune priming (i.e. vaginal birth, daycare attendance, and high birth order)[1] and having fewer infections requiring medical treatment[2] have been inversely associated with disease, suggesting an etiological role of infectious agents, perhaps *via* dysregulation of the immune system.

Patterns of somatic mutation in ALL tumors give further insight into disease etiology. An innate immune enzyme, APOBEC3B (Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3B), inhibits viral infection by inducing TpC>T nucleotide changes in foreign nucleic acid. This mutation signature has been identified in tumor genomes of several cancer types with known (cervical, head and neck, stomach[3]) or hypothesized (breast[4]) infectious etiologies, and is attributed to aberrant enzymatic activity of APOBEC3B. The APOBEC3B point mutation signature is also predominant in ALL,[5] but with subtype specificity. The signature is present in *ETV6-RUNX1* fusion ALL[6] but absent in high hyperdiploid ALL.[7] The high expression of APOBEC3B in lymphoblasts further justifies examination of functional *APOBEC3B* polymorphisms in ALL etiology.

An ~30kb germline deletion polymorphism at the *APOBEC3B* locus has been associated with increased risk of several cancers that bear the APOBEC3B mutation signature[8] and studies have shown that the deletion transcript yields an enzyme with a higher *in vitro* propensity for collateral genomic DNA damage than its wild-type counterpart.[9] In fact, the signature TpC>T point mutation burden is higher in the tumors of ALL and breast cancer patients carrying the germline *APOBEC3B* deletion compared with those without.[9] The deletion is common in populations of Native American ancestry (approx. 60%), and relatively rare in Europeans and Africans (6% and 0.9%, respectively).[10] Hispanic children, whose genetic ancestry is typically comprised of a mixture of Native American, European, and African ancestries, are at the greatest risk for developing ALL in the United States.[11] While it has been suggested that the *APOBEC3B* deletion polymorphism contributes to the patterns of somatic mutations observed in ALL tumors,[9] it is not known whether the variant contributes to disease risk. Here, we report results from the first association study of germline *APOBEC3B* variants in childhood ALL risk.

The *APOBEC3B* deletion genotype was assessed in California Childhood Leukemia Study (CCLS) case and control subjects (see *Online Supplementary Methods* for enrollment details) with a polymerase chain reaction (PCR)-based assay (n=1126). The deletion was tested for association with childhood ALL status overall and within ETV6-RUNX1 fusion and high hyperdiploid ALL subtypes. Overall, controls tended to be wealthier than cases with a higher proportion self-reporting as white and non-Hispanic (Table 1).

APOBEC3B deletion copy number was detected using a validated polymerase chain reaction (PCR) method described previously.[10] A total of 518 ALL cases and 608 controls were genotyped using this PCR assay, with copy number (homozygote wild-type, heterozygote, and homozygote deletion) determined from agarose gel electrophoresis results (*Online Supplementary Figure S1*). A χ^2 test for Hardy-Weinberg equilibrium was performed; the null hypothesis of deletion equilibrium was accepted among controls after stratifying by Hispanic *versus* non-Hispanic ethnicity (P=0.45 and 0.45, respectively). Ethnic heterogeneity in the CCLS population is supported by the distribution of multidimensional scaling (MDS) components compared to reference populations (*Online Supplementary Figure S2*).

After adjusting for global genetic ancestry (first 3 MDS components), no association was observed between the *APOBEC3B* deletion and overall ALL risk for the additive, dominant, or recessive models, nor after stratification by cytogenetic subtype (Table 2). When study subjects were stratified by self-reported Hispanic status, results did not change (*Online Supplementary Table S1*).

Previous studies have identified SNPs within the *APOBEC3* region that are associated with cancer risk independent of the *APOBEC3B* deletion.[8,12] Thus, we tested all SNPs across the *APOBEC3* gene region (chr22:39,200,000-39,650,000) that passed quality filtering (8275 SNPs) for association with ALL in 1083 cases and 1137 controls. After controlling for genetic ancestry, no variant reached statistical significance after correcting for multiple testing (*Online Supplementary Figure S3*). A SNP ~20Kb upstream of *APOBEC3B* and previously associated with bladder cancer,[8] rs1014971, was not associated with ALL risk [Odds Ratio (OR) 0.91, P=0.33]. The top association was seen for rs73424730 (OR 1.35, P=0.004), a SNP ~ 100Kb downstream of the *APOBEC3H* gene.

Table 1. CCLS study participants' characteristics.

	Cases (n=1083)		Controls (n=1137)	
	N	%	N	%
Genome-wide SNP genotyped	1080	99.7	901	79.2
PCR genotyped	518	47.8	608	53.5
Interviewed	975	90.0	1137	100.0
Mean age at Dx				
(Range)	5.54(0.0-14.96)		5.49(0.0-14.93)	
Missing	124	11.4	0	0
Sex				
Male	621	57.3	652	57.3
Female	462	42.7	485	42.7
Missing	0	0	0	0
Ethnicity				
Hispanic	498	46.0	515	45.3
Non-Hispanic	461	42.6	622	54.7
Missing	124	11.4	0	0
Maternal race				
White/Caucasian	787	72.7	983	86.5
African American	34	3.1	34	3.0
Mixed/Other	138	12.7	120	10.6
Missing	124	11.4	0	0
Household income				
<15,000	154	14.2	109	9.6
15,000-29,999	187	17.3	150	13.2
30,000-44,999	143	13.2	141	12.4
45,000-59,999	133	12.3	162	14.2
60,000-74,999	59	5.4	125	11.0
≥75,000	283	26.1	450	39.6
Missing	124	11.4	0	0

CCLS: California Childhood Leukemia Study; SNP: single nucleotide polymorphism; PCR: polymerase chain reaction; Dx: diagnosis.

Table 2. Odds ratios for the association between the APOBEC3B deletion polymorphism and risk of childhood acute lymphoblastic leukemia (ALL) overall and stratified by cytogenetic subtype.

	Cases (n)	Controls (n)	Model	OR*	95% CI	P
Overall	518	608	Additive	0.96	0.77-1.22	0.77
wt/wt	360	441	–	ref	–	–
wt/del	143	144	Dominant	1.03	0.79-1.36	0.81
del/del	15	23	Recessive	0.62	0.31-1.21	0.16
Common ALL	146	608				
wt/wt	100	441	–	ref	–	–
wt/del;del/del	46	167	Dominant	1.08	0.71-1.62	0.73
Hyperdiploid	117	608				
wt/wt	80	441	–	ref	–	–
wt/del;del/del	37	167	Dominant	0.84	0.54-1.34	0.47
t1221	64	608				
wt/wt	47	441	–	ref	–	–
wt/del;del/del	17	167	Dominant	1.10	0.60-2.07	0.77

n: number; ref: reference value; OR: Odds Ratio; CI: Confidence Interval; wt: wild-type; wt: wild-type; del: APOBEC3B deletion; ref: reference. *Adjusted for global genetic ancestry and sex.

To determine whether the observed absence of association between the *APOBEC3B* deletion and ALL was due to confounding by genetic ancestry, local ancestry at the *APOBEC3* megalocus was inferred using a discriminative modeling approach, RFMix.[13] The *APOBEC3B* deletion polymorphism is a highly population-stratified genetic variant.[10] Thus, in the admixed CCLS population, it is possible that adjustment for global genetic ancestry was insufficient, resulting in residual confounding or a washing-out of the true effect of interest. After adjusting for regional ancestral proportions for four continental ancestries, there remained no association between the *APOBEC3B* deletion variant and childhood ALL overall (*Online Supplementary Table S2*). To ensure that effect heterogeneity was not washing-out true associations, RFmix-assigned genetic ancestries (African, Native American, European, and East Asian) at each SNP in the *APOBEC3* locus were tested for independent association with ALL to determine whether local genetic ancestry was associated with disease. African ancestry at the *APOBEC3* megalocus was nominally associated with ALL risk, but did not reach statistical significance after correcting for multiple tests (*Online Supplementary Figure S4*).

Despite the previously observed presence of the *APOBEC3B*-mediated point mutation signature in the tumor genomes of a subset of ALL patients, and a higher burden of point mutations among *APOBEC3B* deletion carriers, we found no apparent relationship between the germline *APOBEC3B* deletion and risk of developing the disease. Moreover, we found no evidence of association between any SNPs across the *APOBEC3* gene region and childhood ALL, including at a locus previously associated with the APOBEC mutation signature in bladder cancer.[14]

Evidence from studies of other cancers suggests a complex relationship between germline variation at the *APOBEC3* gene region and tumorigenesis. For instance, the deletion is associated with an increased risk of breast cancer[9] but appears protective in bladder cancer.[8] Further, the precise contribution of germline *APOBEC* polymorphisms to the presence of the APOBEC mutation signature in an individual tumor has yet to be determined.[15] There is evidence that the *APOBEC3B* deletion polymorphism changes the mutagenic behavior of the enzyme;[9] however, the presence of the mutational signature in a tumor genome does not imply that the deletion is present, as the signature could arise by some other means (i.e. increased enzyme expression by other mechanisms).

Childhood ALL is a rare and heterogeneous disease, and more prevalent in admixed populations, making the study of any potential germline genetic risk factors challenging. While it is unlikely that the *APOBEC3B* deletion polymorphism is a strong independent risk factor for disease overall, our sample size was limited in statistical power to assess small effects, especially following stratification by cytogenetic subtype. Associations observed in other cancers suggest the deletion alters disease risk by 20-30%.[8] These cancers have mixed somatic mutation signatures, suggesting mutagenesis occurs from multiple sources.[5] Conversely, the APOBEC mutational signature is the only one observed in ALL other than common, spontaneous cytosine deamination;[5] therefore, we hypothesized a moderate association with the *APOBEC3B* deletion would be present in ALL. Further, the CCLS study population is representative of cases occurring in the state of California, and reflects the substantial racial and ethnic diversity to be found there. Thus, residual confounding by genetic admixture remains a challenge in investigating this variant, which differs significantly in frequency across populations based on ancestral origin. However, stratifying by self-reported Hispanic status in this study yielded no apparent association in either group. To further account for potential confounding effects of admixture, local ancestry at the *APOBEC3* megalocus was estimated. Adjusting for local ancestral proportions did not change the observed null association of the *APOBEC3B* deletion polymorphism with ALL risk, nor were any of the four observed regional genetic ancestries associated independently with ALL risk.

The apparent lack of association between the germline deletion of *APOBEC3B* and risk of ALL does not provide support that this mutagen, active in tumor cells, is a driver of tumorigenesis. In a previous study, our group showed expression of double-stranded (ds)DNA viruses in primary, treatment naïve, childhood ALL tumors.[16] dsDNA viruses are a primary target of the APOBEC3B enzyme. Transient expression of these viruses in leukemic cells could thus pro-

duce a passenger-type APOBEC3B point mutation signature in ALL. A virus-mediated induction of APOBEC3B expression may be dominant over the impact on APOBEC3B expression by the polymorphism studied here. Alternatively, the APOBEC mutational signature may reflect an infectious etiology of ALL in a subset of cases with heritable predisposition occurring in unrelated genes. Experimental studies in mice have shown that PAX5 mutation, a predisposing risk factor in some childhood ALL cases, can result in development of ALL following exposure to common infection.[17] Though there is no apparent relationship between the germline *APOBEC3B* deletion polymorphism and ALL risk, delineating a potential role for the polymorphism in tumor progression may have treatment implications, warranting further study.

Amelia D. Wallace,[1] Stephen S. Francis,[2,3] Xiaorong Shao,[1] Adam J. de Smith,[2] Kyle M. Walsh,[2] Roberta Mckean-Cowdin,[4] Xiaomei Ma,[5] Gary Dahl,[6] Lisa F. Barcellos,[1] Joseph L. Wiemels[5] and Catherine Metayer[1]

[1]School of Public Health, University of California, Berkeley, CA; [2]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA; [3]Division of Health Sciences, University of Nevada, Reno, NV; [4]Preventative Medicine, University of Southern California, Los Angeles, CA; [5]Epidemiology, Yale, New Haven, CN and [6]Pediatrics Hematology/Oncology, Stanford University, Palo Alto, CA, USA

Acknowledgments: the authors would like to thank the families for their participation. We also thank the clinical investigators at the following collaborating hospitals for help in recruiting patients: University of California Davis Medical Center (Dr. Jonathan Ducore), University of California San Francisco (Drs. Mignon Loh and Katherine Matthay), Children's Hospital of Central California (Dr. Vonda Crouse), Lucile Packard Children's Hospital (Dr. Gary Dahl), Children's Hospital Oakland (Dr. James Feusner), Kaiser Permanente Roseville (former Sacramento) (Drs. Kent Jolly and Vincent Kiley), Kaiser Permanente Santa Clara (Drs. Carolyn Russo, Alan Wong and Denah Taggar), Kaiser Permanente San Francisco (Dr. Kenneth Leung) and Kaiser Permanente Oakland (Drs. Daniel Kronish and Stacy Month). Finally, we acknowledge late Professor Patricia Buffler, who was the Principal Investigator of the CCLS. We also thank the entire California Childhood Leukemia Study staff and the former UCB Survey Research Center for their effort and dedication.

Correspondence: adw222@berkeley.edu
doi:10.3324/haematol.2017.179317

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Ma X, Buffler PA, Selvin S, et al. Daycare attendance and risk of childhood acute lymphoblastic leukaemia. Br J Cancer. 2002;86(9):1419-1424.
2. Chang JS, Tsai CR, Tsai YW, Wiemels JL. Medically diagnosed infections and risk of childhood leukaemia: a population-based case-control study. Int J Epidemiol. 2012;41(4):1050-1059.
3. de Martel C, Ferlay J, Franceschi S, et al. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. Lancet Oncol. 2012;13(6):607-615.
4. Lawson JS, Heng B. Viruses and breast cancer. Cancers (Basel). 2010;2(2):752-772.
5. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415-421.
6. Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. Nat Genet. 2014;46(2):116-125.
7. Paulsson K, Lilljebjorn H, Biloglav A, et al. The genomic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. Nat Genet. 2015;47(6):672-676.
8. Middlebrooks CD, Banday AR, Matsuda K, et al. Association of germline variants in the APOBEC3 region with cancer risk and enrichment with APOBEC-signature mutations in tumors. Nat Genet. 2016;48(11):1330-1338.
9. Nik-Zainal S, Wedge DC, Alexandrov LB, et al. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. Nat Genet. 2014;46(5):487-491.
10. Kidd JM, Newman TL, Tuzun E, Kaul R, Eichler EE. Population stratification of a common APOBEC gene deletion polymorphism. PLoS Genet. 2007;3(4):e63.
11. Barrington-Trimis JL, Cockburn M, Metayer C, Gauderman WJ, Wiemels J, McKean-Cowdin R. Rising rates of acute lymphoblastic leukemia in Hispanic children: trends in incidence from 1992 to 2011. Blood. 2015;125(19):3033-3034.
12. Gohler S, Da Silva Filho MI, Johansson R, et al. Impact of functional germline variants and a deletion polymorphism in APOBEC3A and APOBEC3B on breast cancer risk and survival in a Swedish study population. J Cancer Res Clin Oncol. 2016;142(1):273-276.
13. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. Am J Hum Genet. 2013;93(2):278-288.
14. Rothman N, Garcia-Closas M, Chatterjee N, et al. A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. Nat Genet. 2010;42(11):978-984.
15. Chan K, Roberts SA, Klimczak LJ, et al. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat Genet. 2015;47(9):1067-1072.
16. Francis SS, Wallace AD, Wendt GA, et al. In utero cytomegalovirus infection and development of childhood acute lymphoblastic leukemia. Blood. 2017;129(12):1680-1684.
17. Martin-Lorenzo A, Hauer J, Vicente-Duenas C, et al. Infection Exposure is a Causal Factor in B-cell Precursor Acute Lymphoblastic Leukemia as a Result of Pax5-Inherited Susceptibility. Cancer Discov. 2015;5(12):1328-1343.