Published in final edited form as:
Cancer Epidemiol Biomarkers Prev. 2024 September 03; 33(9): 1248–1252. doi:10.1158/1055-9965.EPI-24-0189.

Folate metabolism and risk of childhood acute lymphoblastic leukemia: a genetic pathway analysis from the Childhood Cancer and Leukemia International Consortium

Catherine Metayer[1], **Logan G. Spector**[2], **Michael E. Scheurer**[3], **Soyoung Jeon**[4], **Rodney J. Scott**[5,6], **Masatoshi Takagi**[7], **Jacqueline Clavel**[8,9], **Atsushi Manabe**[10], **Xiaomei Ma**[11], **Elleni M. Hailu**[1], **Philip J. Lupo**[3], **Kevin Y. Urayama**[12], **Audrey Bonaventure**[8], **Motohiro Kato**[13], **Aline Meirhaeghe**[14], **Charleston W.K. Chiang**[4], **Libby M. Morimoto**[1], **Joseph L. Wiemels**[4]

[1]Division of Epidemiology, School of Public Health, University of California, Berkeley, California, USA

[2]Division of Epidemiology and Clinical Research, Department of Pediatrics, University of Minnesota, Minnesota, USA

[3]Division of Hematology-Oncology, Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA Texas Children's Cancer and Hematology Centers, Texas Children's Hospital, Houston, Texas, USA

[4]Center for Genetic Epidemiology, Department of Population and Public Health Sciences, University of Southern California, Los Angeles, California USA

[5]Faculty of Medicine and Health, School of Biomedical Science and Pharmacy, Hunter Medical Research Institute, University of Newcastle, New Lambton, New South Wales 2305, Australia

[6]Division of Molecular Medicine, NSW Health Pathology, John Hunter Hospital, Newcastle, New South Wales 2305, Australia

[7]Department of Pediatrics and Developmental Biology, Tokyo Medical and Dental University, Tokyo

[8]Inserm UMR1153, Epidemiology of Childhood and Adolescent Cancers (EPICEA) team, Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), Paris, France

[9]French National Registry of Childhood Cancers, RNHE, Hôpital Paul Brousse, Groupe Hospitalier Universitaire Paris-Sud, AP-HP, Villejuif, and RNTSE, CHRU de Nancy, Vandœuvre-lès-Nancy, France

[10]Department of Pediatrics, Hokkaido University, Sapporo

[11]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut, USA

Corresponding author: Catherine Metayer, MD, PhD, 1995 University Avenue, Suite 265, Berkeley, CA 94704, cmetayer@berkeley.edu.

Conflict of Interest: The authors declare no potential conflicts of interest.

[12]Department of Social Medicine, National Center for Child Health and Development, Tokyo, Japan; Graduate School of Public Health, St. Luke's International University, Tokyo

[13]Department of Pediatrics, The University of Tokyo, Tokyo

[14]Université de Lille, INSERM, Centre Hospitalier Universitaire de Lille, Institut Pasteur de Lille, U1167 - RID-AGE - Facteurs de risque et déterminants moléculaires des maladies liées au vieillissement, F-59000 Lille, France

Abstract

Background: Prenatal folate supplementation has been consistently associated with a reduced risk of childhood lymphoblastic leukemia (ALL). Previous germline genetic studies examining the one carbon (folate) metabolism pathway were limited in sample size, scope, and population diversity, and led to inconclusive results.

Methods: We evaluated whether ~2,900 single nucleotide polymorphisms (SNPs) within 46 candidate genes involved in the folate metabolism pathway influence the risk of childhood ALL, using genome-wide data from nine case-control-studies in the Childhood Cancer and Leukemia International Consortium (n=9,058 cases including 4,510 children of European ancestry, 3,018 Latinx, and 1,406 Asians, and 92,364 controls). Each study followed a standardized protocol for quality control and imputation of genome-wide data, and summary statistics were meta-analyzed for all children combined and by major ancestry group using METAL software.

Results: None of the selected SNPs reached statistical significance, overall and for major ancestry groups (using adjusted Bonferroni p-value of 5×10^{-6} and less stringent p-value of 3.5×10^{-5} accounting for the number of "independent" SNPs). None of the 10 top (non-significant) SNPs and corresponding genes overlapped across ancestry groups.

Conclusion: This large meta-analysis of original data does not reveal associations between many common genetic variants in the folate metabolism pathway and childhood ALL in various ancestry groups.

Impact: Genetic variants in the folate pathway alone do not appear to substantially influence childhood ALL risk. Other mechanisms such as gene-folate interaction, DNA methylation or maternal genetic effects may explain the observed associations with self-reported prenatal folate intake.

Keywords

Childhood leukemia; genes; folate; meta-analysis

INTRODUCTION

Leukemia is the most common cancer in children, comprised primarily of acute lymphoblastic leukemia (ALL). One-carbon micronutrients such as folic acid play an essential role in the maintenance of genomic integrity and epigenetic control. Pooled analyses of original data from the Childhood Cancer and Leukemia International Consortium (CLIC) have shown that self-reported prenatal folate and vitamin supplementation reduces the risk of childhood ALL (1). However, germline genetic studies investigating the role of the one carbon (folate) metabolism and childhood ALL risk mostly in European populations have been limited in size and scope focusing on single genes such as *MTHFR*, *TS*, *MTR*, and *MTRR*, and generally yielding inconsistent results (2). We conducted a meta-analysis of CLIC genetic data to investigate the role of ~2,900 candidate single nucleotide polymorphisms (SNPs) in the folate metabolism pathway among diverse populations.

MATERIALS AND METHODS

This study is based on genome-wide data from nine CLIC case-control studies in Europe, North America, Asia, and Oceania, including 9,058 childhood ALL cases and 92,364 study-specific and publicly available controls (Table 1). Each study was given standardized quality control (QC) guidelines for generating genome-wide data, as following: 1) pre-imputation QC (separately for cases and controls if genotyped separately) included filters for SNP call rate <98%, sample call-rate per person <95%, Hardy Weinberg Equilibrium p<10–5 in controls, minor allele frequency (MAF) <0.01; genome-wide identity by descent (IBD) > 0.20, and genome heterozygosity rate within 6sd of mean; 2) for populations with multiple ancestries, principal component analysis (PCA) was performed with known ancestral populations to identify racial and ethnic groups (Europeans, Asians, Latinx, and Black individuals), and exclude population outliers; 3) PCAs were generated on post QC data for adjustment in association analyses; 4) missing data were imputed to HRC reference panel, and 5) post-imputation QC thresholds included MAF<0.01 and $r^2<0.5$. Each study conducted their analyses independently, separately by race and ethnicity (if applicable) using SNPTEST or Plink2, adjusting for PC eigenvectors as appropriate. Prior to sharing summary statistics, each study was asked to assess for genomic inflation and adjust accordingly (lambda<1.1 was considered sufficient). Summary results for each study, including snpID (chr:position), alleles, allele frequency, risk estimate, standard error, p-value, genome build, separately by race/ethnicity, were uploaded to a secure portal. Details on each study are published elsewhere (3–8).

We identified 46 genes in the folate metabolism pathway by curating biological pathways in Gene Ontology, KEGG, GSEA/MSigDB (Broad Institute), USC Genome Browser, and Bioconductor (R) databases, and by reviewing published literature (Table 2). Each selected gene was annotated from the Genome Assembly GRCh37/hg19 using the Bioconductor R package, and SNPs were extracted within 5 kb up- and down-stream from each gene location using UCSC genome table browser, leading to 7,979 candidate SNPs. Genome-wide meta-analyses were conducted using METAL software (version March 2011) for 9,058 ALL cases combined and for the major ancestry subgroups separately i.e., European (n=4,510 cases), Latinx (n=3,018 cases), and Asian (n=1,406 cases). SNPs were included in the meta-analysis if (1) they were available in at least two studies and among > 50,000 subjects overall or of European ancestry and > 10,000 subjects of Asian or Latinx ancestry, and (2) the allele frequency difference across studies was <0.5 among controls (as a quality control check), resulting in ~2,900 SNPs available for analysis (total and European [n=2,855], Latinx [n=2,930], Asian [n-2,230]).. To account for multiple testing, we applied Bonferroni correction (adjusted p-value = 5×10^{-6}) and a less stringent correction defined by the number of “independent” SNPs (based upon 1000 Genomes, calculating the pairwise genotypic correlation using a 100-SNP window, a 10-SNP shift, and a r^2 threshold of 0.2, which average to 350 independent SNPs) and the number of test for each 4 group examined (total, and Europeans, Latinx, and Asian ancestries) resulting in an adjusted p-value of 3.5×10^{-5} (0.05/350/4).

The study was approved by Institutional Review Boards for the California Health and Human Services and the University of California, Berkeley, and was conducted according to the U.S Common Rule.

Data availability:

Only summary statistics were shared by participating studies and no new data were generated as part of this analysis. Original study specific data may be available at the discretion of the individual study principal investigators (information may be requested from the corresponding author).

RESULTS

None of the selected SNPs in the folate metabolism pathway reached the levels of significance defined above, overall and for the three major ancestry groups. Table 3 presents the top 10 SNPs for all groups combined and by ancestry, with crude p-values. None of the 10 top SNPs (and corresponding genes) in each ancestry group overlapped (i.e., *C2orf83, MTHFD1L, NXPH4, SHMT2*, and *SLC19A3* in Europeans; *CBS, GCH1*, and *LRP2* in Latinx; and *CTH, FOLH1* and *NOX4* in Asians).

DISCUSSION

This CLIC study is the largest and most comprehensive to date to investigate the role of genetic variants in the folate metabolism pathway and childhood ALL risk among populations of diverse ancestries. We did not observe statistically significant associations with ~2,900 SNPs. Inherited genetic variants in the folate pathway alone do not appear to substantially influence childhood ALL risk. Alternatively, gene-folate interaction, epigenetic mechanisms, or maternal genetic effects may contribute to the risk.

ACKNOWLEDGMENTS:

Funding:

This study was funded by the UK Children with Cancer grant # 19-308 (C. Metayer, L.M. Morimoto, E. Hailu, J.L. Wiemels). Funding for acquisition of original data in each participating study is listed below by alphabetical order:

Cancer Prevention and Research Institute of Texas RP160771 and RP210064 (M.E. Schuerer)

The collection of samples and data from the patients with childhood ALL was funded by the National Health and Medical Research Council of Australia (https://www.rgms.nhmrc.gov.au) Grant number APP254534. Genotyping was funded by the Hunter Medical Research Institute - the Lawrie Bequest Paediatric Oncology Grant, and the Hunter Children's Research Foundation.

The National Institute of Health grants # P42ES004705 (C. Metayer), R01ES009137 (C. Metayer, L.M. Morimoto), and R24ES028524 (C. Metayer, L.M. Morimoto). Biospecimens and/or data used in this study were obtained from the California Biobank Program, (CBP requests #26 and #1531), Section 6555(b), 17 CCR. The California Department of Public Health is not responsible for the results or conclusions drawn by the authors of this publication.

National Institutes of Health grants # R01CA155461 (J.L. Wiemels and X. Ma).The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement U58DP003862–01 awarded to the California Department of Public Health. This study makes use of data generated by the Wellcome Trust Case–Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113 and 085475. Data came from a grant, the Resource for Genetic Epidemiology Research in Adult Health and Aging (RC2 AG033067; Schaefer and Risch, PIs) awarded to the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH) and the UCSF Institute for Human Genetics. The RPGEH was supported by grants from the Robert Wood Johnson Foundation, the Wayne and Gladys Valley Foundation, the Ellison Medical Foundation, Kaiser Permanente Northern California, and the Kaiser Permanente National and Northern California Community Benefit Programs. The RPGEH and the Resource for Genetic Epidemiology Research in Adult Health and Aging are described here: https://divisionofresearch.kaiserpermanente.org/genetics/rpgeh/rpgehhome. Biospecimens and/or data used in this study were obtained from the California Biobank Program, (CBP request #1380), Section 6555(b), 17 CCR. The California Department of Public Health is not responsible for the results or conclusions drawn by the authors of this publication.

dbGAP accession number: phs000638.v1.p.

Fondation de France, Fondation ARC pour la recherche sur le cancer, AFSSAPS, Cent pour Sang la Vie, Inserm, AFSSET, ANR, Institut National du Cancer INCa, Cancéropôle Ile de France and the Agence Française de Sécurité Sanitaire du Médicament et des Produits de Santé (ANSM);

INCa, Ligue Nationale contre le Cancer, association Enfants et Santé, ANSES, the Agence Nationale de Sécurité Sanitaire de l'alimentation, de l'Environnement et du Travail (PNREST Anses, Cancer TMOI AVIESAN, 2013/1/248), INCa-DHOS, Cancéropôle Ile de France, ANR (Grant id : ANR-10-COHO-0009), Fondation ARC pour la recherche sur le cancer. We are grateful to the SFCE (Société Française de lutte contre les Cancers et leucémies de l'Enfant et l'adolescent), all paediatric oncologists and biologists involved in the two studies, Claire Mulot and the CRB Epigenetec, the Fondation Jean Dausset-CEPH (Centre d'Etude du Polymorphisme Humain), the CEA/CNRGH (Centre National de Recherche en Génomique Humaine), and the SU.VI.Max and the MONALISA Lille studies.

St. Luke's Life Science Institute (Tokyo, Japan), Japan Society for the Promotion of Science (JSPS) KAKENHI grant number 26253041), Japan Agency for Medical Research and Development (grant numbers 15km0305013h0101, 16km0405107h0004, 21kk0305014), the Children's Cancer Association of Japan, and the Japan Leukemia Research Fund.

Abbreviations:

ALL	Acute lymphoblastic leukemia
CLIC	Childhood Cancer and Leukemia International Consortium

REFERENCES

1. Metayer C, Milne E, Dockerty JD, Clavel J, Pombo-de-Oliveira MS, Wesseling C, et al. Maternal supplementation with folic acid and other vitamins and risk of leukemia in offspring: a Childhood Leukemia International Consortium study. Epidemiology 2014;25(6):811–22 doi 10.1097/EDE.0000000000000141. [PubMed: 25207954]
2. Cantarella CD, Ragusa D, Giammanco M, Tosi S. Folate deficiency as predisposing factor for childhood leukaemia: a review of the literature. Genes Nutr 2017;12:14 doi 10.1186/s12263-017-0560-8. [PubMed: 28588742]
3. Ajrouche R, Chandab G, Petit A, Strullu M, Nelken B, Plat G, et al. Allergies, genetic polymorphisms of Th2 interleukins, and childhood acute lymphoblastic leukemia: The ESTELLE study. Pediatric blood & cancer 2022;69(3):e29402 doi 10.1002/pbc.29402. [PubMed: 34662484]

4. Hangai M, Kawaguchi T, Takagi M, Matsuo K, Jeon S, Chiang CWK, et al. Genome-wide assessment of genetic risk loci for childhood acute lymphoblastic leukemia in Japanese patients. Haematologica 2023 doi 10.3324/haematol.2023.282914.
5. Hungate EA, Vora SR, Gamazon ER, Moriyama T, Best T, Hulur I, et al. A variant at 9p21.3 functionally implicates CDKN2B in paediatric B-cell precursor acute lymphoblastic leukaemia aetiology. Nat Commun 2016;7:10635 doi 10.1038/ncomms10635. [PubMed: 26868379]
6. Kennedy AE, Kamdar KY, Lupo PJ, Okcu MF, Scheurer ME, Dorak MT. Genetic markers in a multi-ethnic sample for childhood acute lymphoblastic leukemia risk. Leukemia & lymphoma 2015;56(1):169–74 doi 10.3109/10428194.2014.910662. [PubMed: 24707947]
7. 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(12):2561–4 doi 10.1038/leu.2012.148. [PubMed: 22660188]
8. Wiemels JL, Walsh KM, de Smith AJ, Metayer C, Gonseth S, Hansen HM, et al. GWAS in childhood acute lymphoblastic leukemia reveals novel genetic associations at chromosomes 17q12 and 8q24.21. Nat Commun 2018;9(1):286 doi 10.1038/s41467-017-02596-9. [PubMed: 29348612]

Table 1.

Participants by country/study and ancestry: Childhood Cancer and Leukemia International Consortium

Country[1]	Study name (period)	Overall	Cases	Controls
Australia	Aus-ALL (1998–2006)	1,550	358	1,192
France	ESCALE (2003–2004)[5]	1,983	441	1,542[2]
	ESTELLE (2010–2011)[5]	1,758	343	1,415[3]
Japan	TCCSG (1990–2011)	4,254	540	3,714
	JPLSG (2012–2018)	2,149	548	1,601
United States	ACCESS, Texas (2005-ongoing)[5]	6,965	658	6,307
	CCLS, California (1995–2009)	2,011	1,184	827
	CCRLP, California (1988–2011)	76,317	3,482	72,835[4]
	COG, US-wide (2000–2014)	4,435	1,504	2,931
Total				
All combined		101,422	9,058	92,364
Major ancestry groups				
European		74,521	4,510	70,011
Latinx		12,972	3,018	9,954
Asian		11,738	1,406	10,332

Abbreviations: CCLS: California Childhood Leukemia Study; CCRLP: California Childhood Cancer Record Linkage Project, which does not overlap with CCLS); COG: children Oncology Group; JPLSG: Japanese Pediatric Leukemia/Lymphoma Study Group; TCCSG: Tokyo Children Cancer Study Group.

[1] Alphabetical order

[2] Generic controls from the SU.VI.Max study, France

[3] Generic controls from the MONALISA Lille study, France

[4] Includes publicly available controls from the Wellcome Trust Case–Control Consortium and Resource for Genetic Epidemiology Research in Adult Health and Aging awarded to the Kaiser Permanente Research Program on Genes, Environment, and Health and the University of California San Francisco Institute for Human Genetics, United States.

[5] Estimated proportion of B-cell/T-cell for studies with available subtype information: ESCALE (84%/16%), ESTELLE (80%/20%), ACCESS (89%/11%).

Table 2.

Selected genes in the folate metabolism pathway

AHCY	DHFRL1	MPST	RTBDN
ALDH1L1	DPEP1	MTHFD1	SARDH
ALDH1L2	FOLH1	MTHFD1L	SHMT1
AMT	FOLR1	MTHFD2	SHMT2
ATIC	FOLR2	MTHFD2L	SLC19A1
ATPIF1	FOLR3	MTHFR	SLC19A2
BHMT	FPGS	MTHFS	SLC19A3
C2orf83	FTCD	MTR	SLC25A32
CBS	GART	MTRR	SLC46A1
CPS1	GCH1	MUT	TYMS
CTH	GGH	NOX4	
DHFR	LRP2	PIPOX	

Table 3.

Top 10 single nucleotide polymorphisms and corresponding genes, sorted by crude p-value of the meta-risk estimate for all subjects combined and by ancestry group: Childhood Cancer and Leukemia International Consortium.

Rs#	Symbol	Gene	Reference allele frequency	Beta coefficient	P-value
Total					
rs2239910	SLC46A1	solute carrier family 46 (folate transporter), member 1/sterile alpha and TIR motif containing 1	0.3643	0.0788	2.65E-04
rs9371202	MTHFD1L	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	0.8455	0.1103	4.35E-04
rs12947270	SLC46A1	solute carrier family 46 (folate transporter), member 1/H3 histone, family 3B (H3.3B) pseudogene 2	0.675	−0.0781	5.28E-04
rs9322291	MTHFD1L	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	0.865	0.1397	6.31E-04
rs34449727	CPS1	carbamoyl-phosphate synthase 1, mitochondrial	0.3292	−0.078	7.61E-04
rs11679391	SLC19A3	solute carrier family 19 member 3	0.3726	0.0777	8.36E-04
rs2268369	LRP2	low density lipoprotein receptor-related protein 2	0.5444	−0.0645	1.09E-03
rs2268367	LRP2	low density lipoprotein receptor-related protein 2	0.5445	−0.0643	1.12E-03
rs11886318	LRP2	low density lipoprotein receptor-related protein 2	0.5349	−0.0635	1.34E-03
rs28785011	MTHFD1L	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	0.8654	0.1345	1.40E-03
European					
rs11679391	SLC19A3	solute carrier family 19 (thiamine transporter), member 3	0.4029	0.1107	3.55E-04
rs9371202	MTHFD1L	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	0.8636	0.1576	5.33E-04
rs2138406	C2orf83	chromosome 2 open reading frame 83	0.1873	0.1185	1.21E-03
rs7601819	SLC19A3	solute carrier family 19 (thiamine transporter), member 3	0.8777	0.1626	1.24E-03
rs7583413	C2orf83	chromosome 2 open reading frame 83	0.8086	−0.1156	1.32E-03
rs76758508	SHMT2	serine hydroxymethyltransferase 2	0.315	0.0958	1.63E-03
rs68176600	NXPH4	neurexophilin 4	0.6767	−0.0949	1.69E-03
rs11679339	SLC19A3	solute carrier family 19 (thiamine transporter), member 3	0.7727	−0.1108	1.74E-03
rs4973234	SLC19A3	solute carrier family 19 (thiamine transporter), member 3	0.7727	−0.1093	1.96E-03
rs803456	MTHFD1L	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	0.5117	0.0907	2.24E-03
Latinx					
rs8018688	GCH1	GTP cyclohydrolase 1	0.7902	0.1384	1.38E-03
rs9980564	CBS	cystathionine-beta-synthase	0.6155	0.1202	1.60E-03
rs7147201	GCH1	GTP cyclohydrolase 1	0.7875	0.1298	2.73E-03

Rs#	Symbol	Gene	Reference allele frequency	Beta coefficient	P-value
rs9671455	GCH1	GTP cyclohydrolase 1	0.7462	0.1212	2.78E-03
rs56213135	GCH1	GTP cyclohydrolase 1	0.2014	−0.1308	2.79E-03
rs3759664	GCH1	GTP cyclohydrolase 1	0.1988	−0.13	3.07E-03
rs11886318	LRP2	low density lipoprotein receptor-related protein 2	0.5423	−0.1056	3.24E-03
rs6433109	LRP2	low density lipoprotein receptor-related protein 2	0.5391	−0.1047	3.37E-03
rs7600336	LRP2	low density lipoprotein receptor-related protein 3	0.4182	0.1047	3.73E-03
rs113100590	GCH1	GTP cyclohydrolase 1	0.8052	0.1302	3.74E-03
Asian					
rs11018581	NOX4	NADPH oxidase 4	0.2848	0.2081	7.74E-05
rs11821838	NOX4	NADPH oxidase 4	0.2103	0.196	7.09E-04
rs6677781	CTH	cystathionase	0.2337	0.1782	1.43E-03
rs7925419	FOLH1	folate hydrolase 1	0.4587	0.1463	3.79E-03
rs609054	FOLH1	folate hydrolase 2	0.5818	0.135	6.76E-03
rs2734002	FOLH1	folate hydrolase 3	0.5818	0.1348	6.82E-03
rs10839236	FOLH1	folate hydrolase 4	0.5658	0.1326	8.20E-03
rs3872578	FOLH1	folate hydrolase 5	0.5659	0.1326	8.22E-03
rs9651571	FOLH1	folate hydrolase 6	0.5658	0.1325	8.27E-03
rs7120943	FOLH1	folate hydrolase 7	0.4342	−0.1321	8.44E-03