Replication analysis confirms the association of *ARID5B* with childhood B-cell acute lymphoblastic leukemia

Jasmine Healy,[1] Chantal Richer,[1] Mathieu Bourgey,[1] Ekaterini A. Kritikou,[1] and Daniel Sinnett[1,2]

[1]Division of Hematology-Oncology, Sainte-Justine Hospital Research Center, University of Montreal, QC, Canada, and [2]Department of Pediatrics, University of Montreal, QC, Canada

Citation: Healy J, Richer C, Bourgey M, Kritikou EA, and Sinnett D. Replication analysis confirms the association of ARID5B with childhood B-cell acute lymphoblastic leukemia. Haematologica 2010;95(9):1608-1611. doi:10.3324/haematol.2010.022459

Online Supplementary Table S1. Summary of primers used in the PCR and allele-specific primer extension (ASPE) assays for SNP genotyping.

Gene (Chr)	DNA variant	Position	PCR primers	Product size (bp)	Allele	TAG-ASPE primers
ARID5B (10q21.2)						
	rs7073837	63.369.901	F: ACCAGAATGCACACAGTCTCCTTGC	241	A	TCATTTACCTTTAATCCAATAATCTGGAGAGTGGATCATTTCCCTCA
			R: GGTCCAGCGTGGAAGCCACA		C	TCAATTACCTTTTCAATACAATACGGAGAGTGGATCATTTCCCTCC
	rs10994982	63.380.110	F: ACCTCGTGATCTGCCCGCCT	204	A	CTAATTACTAACATCACTAACAATCATGGTCTTTTAAATATCTTTTGAGAATGCAA
			R: CCACCTCGGCCTTCCGGAGT		G	TCAATTACTTCACTTTAATCCTTTCATGGTCTTTTAAATATCTTTTGAGAATGCAG
	rs10740055	63.388.485	F: ACCACTATGCACTTATCGGAGACAACA	278	G	CAATAAACTATACTTCTTCACTAACACAGGGTTTCTATTTGGAAAGCTGG
			R: GCCCGGCCGTGACCTCTTTT		T	CTTTAATCCTTTATCACTTTATCACACAGGGTTTCTATTTGGAAAGCTGT
	rs10821936	63.393.583	F: ACCAGCTCGGCAGAGCATCC	257	C	TCATAATCTCAACAATCTTTCTTTCTTCTGTGTGCAGTTACTATAGTTGTAC
			R: CCCCGGTGCCTTGAACACACT		T	TCATCAATCAATCTTTTTCACTTTCCTTCTGTGTGCAGTTACTATAGTTGTAT
	rs7089424	63.422.165	F: TGGCTCCCGGTCGTGGCTTA	271	G	TACACTTTCTTTCTTTCTTTCTTTGGTAGCAGTGTTTGGTTATAGTTTAGTTG
			R: CCCAAAACCAAGGGTTTGAACCAGC		T	CAATTTCATCATTCATTCATTTCAGGTAGCAGTGTTTGGTTATAGTTTAGTTT
IKZF1 (7p12.2)						
	rs6964823	50.234.305	F: CTGTCCCTGCAGCGTCTGGC	289	C	TCATCAATCTTTCAATTTACTTACGACGAATGGTCTTGTTCTTCTTCTTC
			R: AGGCCTGGCCGGCTTCTCAG		T	TTACTTCACTTTCTATTTACAATCAGACGAATGGTCTTGTTCTTCTTCTTT
	rs11978267	50.240.513	F: ACTGGCAAGACTGCGGCTGTG	298	C	CAATTCAAATCACAATAATCAATCAGGGGAGGGTAGGTAGAAGTTTATGC
			R: GCTCTCATGGCACGCTCCCC		T	AATCCTTTCTTTAATCTCAAATCACTAGGGGAGGGTAGGTAGAAGTTTATGT
	rs4132601	50.244.813	F: GGGTGTGGCATTTGGAAACGGGA	231	G	TCAAAATCTCAAATACTCAAATCATCACAGAGAAAGATGCGCCTG
			R: ACCAACTGACTCAGGGGGATGGA		T	CAATTAACTACATACAATACATACATCACAGAGAAAGATGCGCCTT
	rs6944602	50.247.960	F: ACACAGCTTCACGGTGTGACCC	184	A	TACACAATCTTTTCATTACATCATCCAGCGGGCATTACACTCA
			R: ATTGCTGCACGGGGGTCTGC		G	ATCATACATACATACAAATCTACACCAGCGGGCATTACACTCG
DDC (7p12.2)						
	rs7809758	50.347.542	F: TGTCCAGCCCAGCCACCTGT	180	C	CTTTATCAATACATACTACAATCAAAGCTGGTGAGACGAACCTTCC
			R: CCCCAGTTCGAGACCTTTGTGGC		T	ATACTACATCATAATCAAACATCAAAAGCTGGTGAGACGAACCTTCT
	rs880028	50.344.345	F: ACTTGGCCCAGGGGACAGCA	106	A	AATCTTACTACAAATCCTTTCTTTGCCCTTTAAATCCGATAGCCCTA
			R: TCCTGGCCAGTAGCTGGCTGAT		G	CTTTTCATCAATAATCTTACCTTTGCCCTTTAAATCCGATAGCCCTG
	rs3779084	50.342.944	F: GGGATGCATGGAGCTGTGGGC	190	A	CTTTTCAAATCAATACTCAACTTTGGCCCATCTAGGAACCAGGCA
			R: CCCCACTGAGGCAGCCTTGC		G	CTACAAACAAACAAACATTATCAAGCCCATCTAGGAACCAGGCG
	rs2242041	50.303.658	F: ACAGCGGTACTTTCCCTCCCCT	99	C	CTATCTTTAAACTACAAATCTAACGTCTTATGCTGAGAGCAATGGAATAAC
			R: CTCAAGCATCTCTGGAAGGTAGTGGGG		G	CTAACTAACAATAATCTAACTAACGTCTTATGCTGAGAGCAATGGAATAAG
CEBPE (14q11.2)						
	rs2239633	22.658.897	F: TGCTGGGCTCCACCTACCCC	159	A	TCATTTCACAATTCAATTACTCAATAGGTCCTAGGAACAAGCTCTACACA
			R: CTCTGGAGCACCACGCAGGC		G	CAATTCATTTACCAATTTACCAATTAGGTCCTAGGAACAAGCTCTACACG
OR2C3 (1q44)						
	rs1881797	244.015.573	TGTGCACAGCATGTGTTGAGATGA	292	C	CTACTATACATCTTACTATACTTTTGGGTTAGTTTCTGAATTCTTGCATTC
			GCACTGGACACAGCTTCTGCCT		T	TTACCTTTATACCTTTCTTTTTACTGGGTTAGTTTCTGAATTCTTGCATTT

SNP position relative to the UCSC Genome Browser Human May 2004 Assembly (hg17)

Online Supplementary Table S2. Replication analysis in the Quebec Childhood ALL cohort of germline SNPs whose allele frequencies differed between children with ALL and control groups in two genome-wide association studies.

Gene (Chr.) DNA variant	Risk allele	QcALL replication study RAF	Subgroup	OR (95% CI)	P value	GWA1 - Papaemmanuil *et al.*[1] RAF	Subgroup	OR (95% CI)	P value	GWA2 - Trevino *et al.*[2] RAF	Subgroup	OR (95% CI)	P value
ARID5B (10q21.2)													
rs7073837	A	0.34	B-cell	1.54 (1.21-1.97)	4.2E-04	0.40	NA	1.58 (1.35-1.89)	4.7E-16			NA	
			B-hyperdip	2.10 (1.51-2.93)	9.6E-06		B-cell	1.59 (1.48-1.71)	1.0E-15				
rs10994982	A	0.48	B-cell	1.55 (1.22-1.97)	3.8E-04					0.47	NA	1.71 (1.43-2.05)	1.2E-09
			B-hyperdip	1.87 (1.34-2.60)	NA 2.0E-04						B-hyperdip	1.71 (1.19-2.46)	0.0025
rs10740055	C	0.49	B-cell	1.71 (1.34-2.19)	1.6E-05	0.50	NA	1.53 (1.41-1.64)	5.4E-14				
			B-hyperdip	2.06 (1.46-2.90)	3.0E-05		B-cell	1.57 (1.45-1.81)	1.6E-14			NA	
rs10821936	C	0.33	B-cell	1.93 (1.51-2.48)	1.7E-07					0.33	NA	2.00 (1.68-2.38)	2.8E-16[a]
			B-hyperdip	2.91 (2.08-4.07)	2.0E-10			NA			B-hyperdip	2.12 (1.49-3.01)	9.7E-06
rs7089424	G	0.43	B-cell	1.91 (1.49-2.45)	3.6E-07	0.34	NA	1.65 (1.54-1.76)	6.7E-19			NA	
			B-hyperdip	2.87 (2.04-4.04)	8.2E-10		B-cell	1.70 (1.58-1.81)	1.4E-19				
							B-hyperdip	NA	3.8E-06[b]				
IKZF1 (7p12.2)													
rs6964823	G	0.54	B-cell	1.15 (0.90-1.46)	0.27	0.50	NA	1.52 (1.41-1.64)	6.0E-14			NA	
							B-cell	1.53 (1.42-1.65)	1.9E-13				
rs11978267	G	0.31	B-cell	1.27 (0.99-1.65)	0.065			NA		0.27	NA	1.69 (1.4-1.9)	8.8E-11
rs4132601	G	0.31	B-cell	1.26 (0.98-1.62)	0.075	0.28	NA	1.69 (1.58-1.81)	1.2E-19[c]			NA	
							B-cell	1.73 (1.61-1.85)	9.3E-20				
rs6944602	A	0.24	B-cell	1.17 (0.89-1.55)	0.27	0.21	NA	1.64 (1.37-2.07)	3.4E-15			NA	
						B-cell		1.69 (1.56-1.81)	1.5E-15				
DDC (7p12.2)													
rs7809758	G	0.41	B-cell	1.02 (0.80-1.30)	0.9	0.37	NA	1.44 (1.32-1.54)	2.4E-10			NA	
							B-cell	1.48 (1.37-1.60)	2.9E-11				
rs880028	C	0.25	B-cell	1.04 (0.79-1.36)	0.79	0.22	NA	1.43 (1.30-1.56)	1.3E-07			NA	
							B-cell	1.49 (1.36-1.61)	1.4E-09				
rs3779084	C	0.25	B-cell	1.01 (0.77-1.33)	0.92	0.22	NA	1.44 (1.32-1.56)	8.8E-09			NA	
							B-cell	1.50 (1.37-1.63)	6.5E-10				
rs2242041	G	0.11	B-cell	1.40 (0.99-2.00)	0.059		NA			0.09	NA	1.72 (1.3-2.1)	9.9E-07
CEBPE (14q11.2)													
rs2239633	G	0.53	B-cell	1.19 (0.94-1.52)	0.15	0.52	NA	1.34 (1.22-1.45)	2.9E-07			NA	
							B-cell	1.37 (1.26-1.49)	5.6E-08				
OR2C3 (1q44)													
rs1881797	C	0.17	B-cell	0.95 (0.66-1.35)	0.77		NA			0.16	NA	1.52 (1.2-1.8)	7.3E-06
			t(12;21)	1.31 (0.53-3.24)	0.57						t(12;21)	2.08 (1.1-3.8)	0.021

Results are shown for allelic case-control association tests; P values in bold remain significant after FDR adjustment for multiple testing at the 5% level. The Quebec Childhood ALL (QcALL) replication cohort consisted of 284 B-cell ALL cases and 270 controls; logistic regression was applied to either the full dataset or to a restricted subgroup of patients as specified. GWA1; the Papaemmanuil et al. study comprised 907 cases (824 B-cell, 83 T-cell) and 2,398 controls; logistic regression was applied to either the full dataset or to a restricted subgroup of patients as specified; P values denote Cochran-Armitage trend test statistics. GWA2; the Trevino et al. study consisted of a discovery cohort of 317 cases (274 B-cell, 43 T-cell) and 17,958 controls; logistic regression was used and subgroup analysis was performed by comparing allele frequencies between single ALL subgroups and all other subgroups combined. Fisher's combined probability method was used to calculate a combined p value of the current QcALL replication study and the previous GWASs. RAF indicates risk allele frequency in controls; OR,odds ratio; CI, confidence interval; B-hyperdip; B-cell hyperdiploid ALL; NA: not applicable. [a]Strongest association signal from the Trevino et al. study. [b]P value denotes case-only logistic regression analysis. [c]Strongest association signal from the Papaemmanuil et al. study.

Online Supplementary Table S3. Distribution of *ARID5B* haplotypes among B-cell ALL cases and controls from the Quebec gender-specific haplotype risks estimates.

Haplotype	DNA variant					B-cell ALL cases	Controls	OR (95% CI)	*P* value	Global χ^2 (df)	Global *P* value
	rs7073837	rs10994982	rs10740055	rs10821936	rs7089424	Total (%)					
ARID5B*1	C	G	A	T	T	224 (38.23)	268 (50.00)	1 (referent)	–	45.03 (14)	4.04E-05
ARID5B*2	A	A	C	C	G	271 (46.25)	168 (31.34)	1.93 (1.47-2.53)	7.6E-07		
ARID5B*3	A	A	C	T	T	31 (5.29)	46 (8.58)	0.81 (0.48-1.35)	0.39		
ARID5B*4	C	A	C	T	T	37 (6.31)	34 (6.34)	1.30 (0.77-2.21)	0.30		
ARID5B*	–	–	–	–	–	23 (3.92)	20 (3.73)	1.38 (0.70-2.71)	0.31		
Males (%)											
ARID5B*1	C	G	A	T	T	120 (34.29)	159 (52.65)	1 (referent)	–	36.95 (16)	2.13E-03
ARID5B*2	A	A	C	C	G	170 (48.57)	92 (30.46)	2.45 (1.70-3.52)	3.41E-07		
ARID5B*3	A	A	C	T	T	19 (5.43)	21 (6.95)	1.20 (0.58-2.46)	0.59		
ARID5B*4	C	A	C	T	T	26 (7.43)	15 (4.97)	2.30 (1.11-4.87)	0.014		
ARID5B*	–	–	–	–	–	15 (4.29)	15 (4.97)	1.32 (0.58-3.03)	0.46		
Females (%)											
ARID5B*1	C	G	A	T	T	104 (44.07)	108 (46.35)	1 (referent)	–	21.45 (9)	0.011
ARID5B*2	A	A	C	C	G	99 (41.95)	75 (32.19)	1.37 (0.90-2.09)	0.120		
ARID5B*3	A	A	C	T	T	13 (5.51)	24 (10.30)	0.56 (0.25-1.22)	0.120		
ARID5B*4	C	A	C	T	T	11 (4.66)	20 (8.58)	0.57 (0.23-1.33)	0.16		
ARID5B*	–	–	–	–	–	9 (3.81)	6 (2.58)	1.56 (0.47-5.50)	0.41		

Logistic regression was used to estimate haplotype-specific ORs in either the full dataset or in restricted subgroups stratified by gender, comparing male cases to male controls or female cases to female controls. The most common haplotype was used as reference. P values in bold remain significant after FDR adjustment for multiple testing at the 5% level. Percentages indicate number of chromosomes with given haplotype/total number of chromosomes. Haplotypes with relative frequencies <5% are grouped under ARID5B and are represented as * combinations of the four DNA variants. A likelihood ratio test was performed in FAMHAP to compare global haplotype differences between cases and controls and is reported here as a Global χ^2 test with number of haplotype parameters different from zero-1 degrees of freedom. OR indicates crude odds ratio; df, degrees of freedom; and —, not applicable. [a]Significant risk difference between males and females based on Mantel-Haenszel χ^2 test of homogeneity: X2= 7.31 (1df); P value= 0.0069.*

References

1. 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. Nat Genet. 2009;41(9):1006-10.
2. Trevino LR, Yang W, French D, Hunger SP, Carroll WL, Devidas M, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet. 2009;41(9):1001-5.