Brief report

Common variation at 6p21.31 (*BAK1*) influences the risk of chronic lymphocytic leukemia

*Susan L. Slager,[1] *Christine F. Skibola,[2] Maria Chiara Di Bernardo,[3] Lucia Conde,[2] Peter Broderick,[3] Shannon K. McDonnell,[1] Lynn R. Goldin,[4] Naomi Croft,[3] Amy Holroyd,[3] Shelley Harris,[3] Jacques Riby,[2] Daniel J. Serie,[1] Neil E. Kay,[1] Timothy G. Call,[1] Paige M. Bracci,[5] Eran Halperin,[6] Mark C. Lanasa,[7] Julie M. Cunningham,[1] Jose F. Leis,[1] Vicki A. Morrison,[8,9] Logan G. Spector,[9] Celine M. Vachon,[1] Tait D. Shanafelt,[1] Sara S. Strom,[10] Nicola J. Camp,[11] J. Brice Weinberg,[7,12] Estella Matutes,[13] Neil E. Caporaso,[4] Rachel Wade,[14] Martin J. S. Dyer,[15] Claire Dearden,[13] James R. Cerhan,[1] Daniel Catovsky,[13] and Richard S. Houlston[3]

[1]Mayo Clinic College of Medicine, Rochester, MN; [2]Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, CA; [3]Molecular and Population Genetics, Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, United Kingdom; [4]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD; [5]Department of Epidemiology and Biostatistics, School of Medicine, University of California, San Francisco, San Francisco, CA; [6]Blavatnik School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel; [7]Duke University Medical Center, Durham, NC; [8]Minneapolis Veterans Administration Medical Center, Minneapolis, MN; [9]University of Minnesota, Minneapolis, MN; [10]The MD Anderson Cancer Center, Houston, TX; [11]University of Utah School of Medicine, Salt Lake City, UT; [12]Durham Veterans Administration Medical Center, Durham, NC; [13]Section of Haemato-Oncology, Division of Pathology, Institute of Cancer Research, Sutton, United Kingdom; [14]Clinical Trial Service Unit, University of Oxford, Oxford, United Kingdom; and [15]Department of Cancer Studies and Molecular Medicine, Leicester University, Leicester, United Kingdom

We performed a meta-analysis of 3 genome-wide association studies to identify additional common variants influencing chronic lymphocytic leukemia (CLL) risk. The discovery phase was composed of genome-wide association study data from 1121 cases and 3745 controls. Replication analysis was performed in 861 cases and 2033 controls. We identified a novel CLL risk locus at 6p21.33 (rs210142; intronic to the *BAK1* gene, BCL2 antagonist killer 1; $P = 9.47 \times 10^{-16}$). A strong relationship between risk genotype and reduced *BAK1* expression was shown in lymphoblastoid cell lines. This finding provides additional support for polygenic inheritance to CLL and provides further insight into the biologic basis of disease development. (*Blood*. 2012;120(4):843-846)

Introduction

Chronic lymphocytic leukemia (CLL) is the most common form of lymphoid malignancy in Western countries.[1] Although CLL shows a strong familial risk,[2] the genetic basis of inherited predisposition to CLL is largely unknown. Recent genome-wide association studies (GWAS) of CLL have provided evidence that the coinheritance of multiple low-risk variants located on chromosomes 2q37.1, 2q37.3, 2q13, 6p21.3, 6p25.3, 8q24.21, 11q24.1, 15q21.3, 15q23, 15q25.2, 16q24.1, and 19q13.32 contributes to the heritability of CLL.[3-6]

The statistical power of individual GWAS has been limited by the modest effect sizes of individual genetic variants, the need to establish stringent statistical significance thresholds, and financial constraints on the number of variants that can be followed up. Meta-analysis of existing GWAS data therefore offers the opportunity to discover additional CLL susceptibility loci.

In this study, we conducted a meta-analysis of GWAS data, followed by validation in an independent case-control series, enabling us to identify a novel susceptibility locus for CLL at 6p21.33.

Methods

Participants

All data collection from study participants were approved by the respective institutional review boards, and all participants provided written informed consent in accordance with the Declaration of Helsinki. For all cases, the diagnosis of CLL had been pathologically confirmed in accordance with the World Health Organization guidelines.[7]

Discovery datasets

The discovery phase was composed of 3 previously described GWAS conducted in the United Kingdom (UK-GWAS)[4,5] with 503 cases and 2699 controls, in the San Francisco Bay Area (SF-GWAS)[8] with 211 cases and 750 controls, and in the Genetic Epidemiology of CLL (GEC) consortium (GEC-GWAS)[3] with 407 cases and 398 controls (supplemental Methods, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).

Replication series

The replication series was composed of 861 CLL cases (565 men; mean age at diagnosis, 61.9 years), ascertained through United Kingdom hematology clinics. Controls were 2033 healthy persons recruited to the National Cancer Research Network genetic epidemiologic studies, the National Study of Colorectal Cancer,[9] the Genetic Lung Cancer Predisposition Study,[10] and the Royal Marsden Hospital Trust/Institute of Cancer Research Family History and DNA Registry. These controls were the spouses or unrelated friends of persons with malignancies. Both cases and controls were British residents and of European ancestry. Genotyping was conducted using competitive allele-specific PCR KASPar chemistry (KBiosciences).

To confirm genotyping accuracy, duplicate samples were genotyped, along with direct sequencing of subsets of samples. For all single nucleotide polymorphisms (SNPs), more than 99% concordant results were obtained.

Statistical analysis

Main analyses used R Version 2.5, Stata Version 11, and PLINK Version 1.07 software.[11] The association between each SNP and CLL risk was assessed by the Cochran-Armitage trend test. Odds ratios (ORs) and 95% confidence intervals were calculated by unconditional logistic regression. Meta-analysis was conducted under a fixed-effects model. Cochran Q statistic (to test for heterogeneity) and the I^2 statistic (to quantify the proportion of the total variation because of heterogeneity) were calculated.[12] Associations by sex, age, and clinic-pathologic phenotypes were examined by logistic regression in case-only analyses. The familial risk attributable to SNPs was calculated as previously described.[13] Linkage disequilibrium metrics were based on Data Release 27/phase 3 (February 2009) on NCBI B36 assembly.

Results and discussion

The combined GWASs provided genotype data on 1121 cases and 3745 controls; imputation based on data from the HapMap project allowed association testing for more than 1 500 000 SNPs. We verified the previously identified risk loci (all $P < .05$) in the meta-analysis (supplemental Table 1) and further identified 15 SNPs that mapped to 6 novel loci.

To validate these findings, we genotyped the top 10 informative SNPs in an additional series of 861 CLL cases and 2033 controls (Table 1; supplemental Table 2).

The 2 SNPs localizing to 6p21.33, rs210134 and rs210142, showed convincing evidence of an association in the replication series (OR = 1.35, $P = 4.87 \times 10^{-6}$ and OR = 1.47, $P = 2.41 \times 10^{-8}$, respectively; Table 1). In the combined analysis, both rs210142 (OR = 1.37, $P = 9.47 \times 10^{-16}$) and rs210134 (OR = 1.40, $P = 1.03 \times 10^{-12}$) provided evidence for an association with CLL at genome-wide significance (ie, $P < 5.0 \times 10^{-8}$; Table 1). None of the other SNPs replicated.

rs210142 maps to chromosome 6p21.33, within intron 1 of the *BAK1* gene (BCL2 antagonist killer 1; MIM: 600516), and rs210134 maps 100 kb telomeric to *BAK1* (supplemental Figure 1). Both SNPs map to the same region of linkage disequilibrium and are highly correlated ($r^2 = 0.8$, $D' = 0.9$). *BAK1* promotes apoptosis by binding to and antagonizing the apoptosis repressor activity of *BCL2* and other antiapoptotic proteins.[14,15] Somatic rearrangements of the immunoglobulin heavy chain locus of *BCL2* that result in constitutive *BCL2* overexpression are found in both CLL and follicular lymphomas. The expression of *BAK1* is essential for the maintenance of B-cell homeostasis; in mice that are conditionally deficient in BAK1, there is an accumulation of immature and mature follicular B cells with defective cell cycling in response to B-cell receptor stimulation.[16]

To explore whether the 6p21.33 association reflects *cis*-acting regulatory effects on *BAK1*, we analyzed publicly available mRNA expression (Figure 1). A strong relationship between rs210134 risk genotype and reduced *BAK1* expression was shown in both datasets ($P = 7.8 \times 10^{-5}$ and .0389, combined $P = 4.4 \times 10^{-5}$; Figure 1). This result suggests a biologically plausible mechanism by which reduction of *BAK1* expression alleviates repression of antiapoptotic proteins, thereby inhibiting apoptosis and hence contributing to B-cell neoplasia. *BAK1* does not, however, directly interact with *BCL2*, but its interaction is dependent on the allelic form of *BCL2* present in cells; thus, the effects of *BAK1* are context dependent.[18]

Table 1. Summary results for rs210134 and rs210142 SNPs associated with CLL risk

SNP	Position, bp	Risk allele	UK-GWAS			SF-GWAS			GEC-GWAS			Replication			Combined		
			RAF	OR (95% CI)	P_{trend}	RAF	OR (95% CI)	P_{trend}	RAF	OR (95% CI)	P_{trend}	RAF	OR (95% CI)	P_{trend}	OR (95% CI)	P_{trend}	P_{het} (I^2)
rs210134	33 648 187	G	0.68	1.31 (1.13-1.51)	2.95×10^{-4}	0.69	1.56 (1.21-2.02)	6.39×10^{-4}	0.70*	1.37 (1.08-1.75)*	.0096*	0.68	1.35 (1.19-1.54)	4.87×10^{-6}	1.37 (1.22-1.53)	1.03×10^{-12}	.50(0%)
rs210142	33 654 815	C	0.70*	1.35 (1.17-1.57)*	6.82×10^{-5}*	0.70*	1.58 (1.22-2.05)*	5.44×10^{-4}*	0.70*	1.38 (1.09-1.76)*	.0077*	0.69	1.47 (1.28-1.68)	2.41×10^{-8}	1.40 (1.25-1.57)	9.47×10^{-16}	.59(0%)

RAF indicates risk allele frequencies (based on control data).
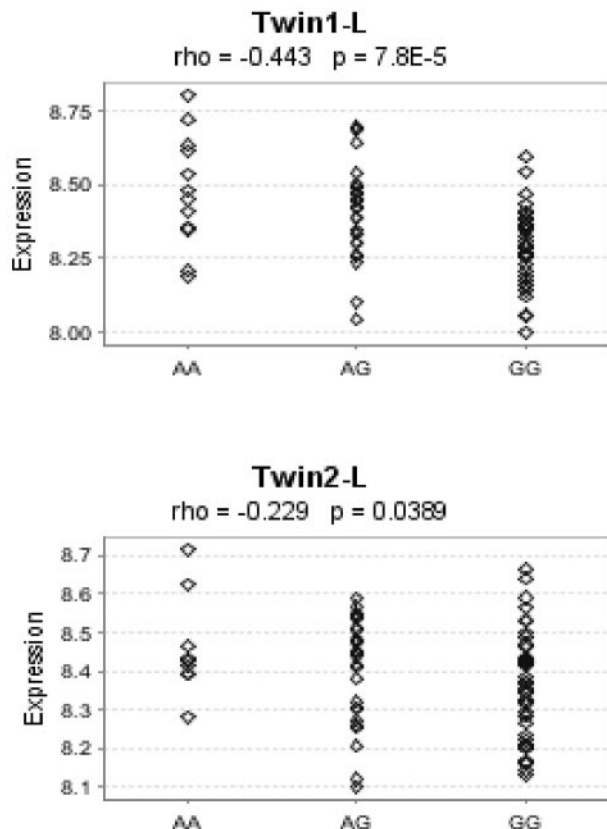
*Imputed genotype.

Figure 1. Lower *BAK1* expression associated with the G allele of rs210134 in lymphoblastoid cell lines. We assessed the correlation between normalized *BAK1* gene expression and the count of the G allele of rs210134 using Spearman rank correlation and publicly available Sentrix Human-6 Expression BeadChips (Illumina) data on 156 lymphoblastoid cell lines derived from healthy female twins as part of the MuTHER project[17] as of December 15, 2008 (GENEVAR project). The G allele is associated with the elevated risk of CLL.

CLL shows male predominance and can be classified on the basis of the presence or absence of somatic hypermutations of *IGVH* genes, with mutated CLL having a better prognosis.[19] We assessed the relationship between age, sex, *IGVH* mutation, and rs210142 and rs210134 genotypes by case-only logistic regression (supplemental Methods). rs210142 and rs210134 showed no evidence of a relationship with age, sex, or *IGVH* mutation status (supplemental Table 3). Furthermore, using data from the CLL4 trial patients (supplemental Methods), we found no evidence that rs210134 genotype influences overall survival or progression-free survival (hazard ratio = 1.11, $P = .41$ and hazard ratio = 1.12, $P = .23$; respectively).

We investigated the combined effect of the 6p21.33 variation and the previously identified risk variants on CLL risk. No evidence of interaction between any of the loci ($P > .05$) was observed, compatible with each locus having an independent role in defining risk. Whereas the risks of CLL associated with the 6p21.33 and other variants are individually modest, the carrier frequencies of the risk alleles are high in those persons of European ancestry; therefore, the loci contribute significantly to the development of CLL.

Collectively, the currently identified susceptibility loci account for approximately 16% of the familial CLL risk. Previous genetic linkage studies have failed to provide evidence that rare, high-penetrance genes contribute significantly to the familial risk. Our study had moderate power (~ 50%) to detect variants, such as

6p21.33, indicating that additional common variants with similar or smaller effects might be identified with additional GWAS data.

By pooling GWAS data and conducting replication analyses, we have identified a novel CLL susceptibility locus. Although additional analyses are required to determine the functional consequences of 6p21.33 variation, the findings further highlight the importance of genetic variation in B-cell developmental pathway as a biologic basis to CLL pathogenesis. The frequency of the *rs210142* risk allele is substantially higher in Europeans than in other ancestry groups consistent with adaptive selection. Hence, it will also be intriguing to explore how our findings translate to non-European populations, some of which are typified by a significantly lower prevalence of CLL.[20]

Authorship

References

1. Stevenson FK, Caligaris-Cappio F. Chronic lymphocytic leukemia: revelations from the B-cell receptor. *Blood.* 2004;103(12):4389-4395.

2. Goldin LR, Pfeiffer RM, Li X, Hemminki K. Familial risk of lymphoproliferative tumors in families of patients with chronic lymphocytic leukemia: results from the Swedish Family-Cancer Database. *Blood.* 2004;104(6):1850-1854.

3. Slager SL, Rabe KG, Achenbach SJ, et al. Genome-wide association study identifies a novel susceptibility locus at 6p21.3 among familial CLL. *Blood.* 2011;117(6):1911-1916.

4. Di Bernardo MC, Crowther-Swanepoel D, Broderick P, et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nat Genet.* 2008;40(10):1204-1210.

5. Crowther-Swanepoel D, Broderick P, Di Bernardo MC, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. *Nat Genet.* 2010;42(2):132-136.

6. Crowther-Swanepoel D, Di Bernardo MC, Jamroziak K, et al. Common genetic variation at 15q25.2 impacts on chronic lymphocytic leukaemia risk. *Br J Haematol.* 2011;154(2):229-233.

7. Swerdlow SH, Campo E, Harris N. *World Health Organization Classification of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon, France: IARC Press; 2008.

8. Conde L, Halperin E, Akers NK, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat Genet.* 2010;42(8):661-664.

9. Penegar S, Wood W, Lubbe S, et al. National study of colorectal cancer genetics. *Br J Cancer.* 2007;97(9):1305-1309.

10. Eisen T, Matakidou A, Houlston R. Identification of low penetrance alleles for lung cancer: the GEnetic Lung CAncer Predisposition Study (GELCAPS). *BMC Cancer.* 2008;8:244.

11. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81(3):559-575.

12. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21(11):1539-1558.

13. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med.* 2008;358(26):2796-2803.

14. Chittenden T, Harrington EA, O'Connor R, et al. Induction of apoptosis by the Bcl-2 homologue Bak. *Nature.* 1995;374(6524):733-736.

15. Leu JI, George DL. Hepatic IGFBP1 is a prosurvival factor that binds to BAK, protects the liver from apoptosis, and antagonizes the proapoptotic actions of p53 at mitochondria. *Genes Dev.* 2007;21(23):3095-3109.

16. Takeuchi O, Fisher J, Suh H, Harada H, Malynn BA, Korsmeyer SJ. Essential role of BAX, BAK in B cell homeostasis and prevention of autoimmune disease. *Proc Natl Acad Sci U S A.* 2005;102(32):11272-11277.

17. Nica AC, Parts L, Glass D, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet.* 2011;7(2):e1002003.

18. Dai H, Meng XW, Lee SH, Schneider PA, Kaufmann SH. Context-dependent Bcl-2/Bak interactions regulate lymphoid cell apoptosis. *J Biol Chem.* 2009;284(27):18311-18322.

19. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood.* 1999;94(6):1848-1854.

20. Weiss NS. Geographical variation in the incidence of the leukemias and lymphomas. *Natl Cancer Inst Monogr.* 1979;(53):139-142.