



Published in final edited form as:

Semin Hematol. 2013 October ; 50(4): . doi:10.1053/j.seminhematol.2013.09.007.

Genetic Susceptibility to Chronic Lymphocytic Leukemia

Susan L. Slager, PhD^{1,*}, Neil E. Caporaso, MD², Silvia de Sanjose, MD^{3,4}, and Lynn R Goldin, PhD²

¹Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota ²Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland ³Unit of Infections and Cancer, Cancer Epidemiology Research Programme, Institut Catala d'Oncologia, IDIBELL, Barcelona, Spain ⁴Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain

Abstract

Chronic lymphocytic leukemia (CLL) is the most common adult leukemia in the West and is an incurable malignancy. No firmly established evidence exists for environmental risk-factors in the etiology of CLL. However, CLL is estimated to have one of the highest familial risks for a hematologic malignancy; this along with other evidence strongly supports an inherited genetic component. In the past five years, genome-wide association studies have provided the foundation for new avenues in the investigation of pathogenesis of this disease with 22 susceptibility loci currently identified. We review here the advances made in identifying these loci, the potential to translate these findings into clinical practice, and future directions needed to advance our understanding of the genetic susceptibility of CLL.

Keywords

CLL; SNP associations; etiology

Introduction

Chronic lymphocytic leukemia (CLL) is a neoplasm of the blood and is one of the most common B-cell malignancies in individuals of European ancestry in the United States; in 2013, ~16,000 individuals are expected to be diagnosed¹. It is clinically defined by the presence of a clonal population of B-cell lymphocytes of more than 5×10^9 cells/L that have a characteristic immunophenotype². CLL is considered a subtype of non-Hodgkin lymphoma (NHL), and because CLL and small lymphocytic leukemia (SLL) have the same immunophenotype and similar clinical course, CLL and SLL are typically combined together, despite the more prominent nodal involvement in SLL. CLL is an adult cancer with

© 2013 Elsevier Inc. All rights reserved.

*Corresponding Author: Susan L. Slager, PhD, Department of Health Sciences Research, Mayo Clinic College of Medicine, 200 1st Street SW, Rochester, MN 55905. Telephone: 507-284-5965; Fax: 507-284-9542; slager@mayo.edu.

Disclosure of Potential Conflicts of Interest: The authors declare that they have no conflicts of interest or competing financial or personal relationships that could inappropriately influence the content of this article.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

a median age at CLL diagnosis ranging between 65–72 years^{3–6}. Between 1975 and 2009, the incidence of CLL was 4.2 per 100,000, increases with age, and was nearly twice as high in males as females⁶. Although there has been a dramatic evolution in the treatment options for the clinical care of CLL patients over the past decade, it still remains an incurable malignancy. It is hoped that expanding our understanding of the pathogenesis of CLL can both suggest translational targets and delay/prevent progression of the precursor to CLL. No environmental risk factor has been consistently and reproducibly identified to be associated with CLL risk⁷. However, CLL has one of the highest familial risks of any cancer, suggesting a strong inherited genetic component⁸. In this review, we present the published studies which demonstrate the strong familial risk of CLL as well as the studies which have identified the currently known 22 CLL susceptibility loci. Furthermore, we discuss both the potential impact these findings could have on clinical practice as well as future directions for research into the genetics of CLL.

Familial risk of CLL

Diverse lines of evidence support a strong inherited genetic component of CLL risk. Initial case-control studies reported a significant association of family history with lymphoproliferative (LP) tumors^{9,10}. More recently, the InterLymph consortium pooled data from 10,211 non-Hodgkin lymphoma (NHL) cases and 11,904 controls to evaluate risk of LP. They found a significant risk of CLL for those with a family history of any LP (1.8 fold increase) or family history of any leukemia (2.2 fold increase risk). In Norway, investigators used a combination of questionnaire data and registry data to examine familial occurrence of CLL. They reported that 18% of CLL patients had a family history of hematologic disease and relatives of patients had a 6-fold risk of developing CLL¹¹. Finally, Goldin *et al.* have published a series of population-based studies linking cancer registry with population-based registries from Sweden^{12,13}. In these analyses, the risk of CLL was evaluated in 26,947 first-degree relatives of 9,717 CLL patients compared to that of 107,223 first-degree relatives of 38,159 matched controls. Relatives of CLL patients had an 8.5 fold increase risk of CLL and a 1.9 fold increase risk of NHL.

Monoclonal B-cell Lymphocytosis (MBL) and CLL families

MBL is an asymptomatic hematologic condition characterized by a small number of B-cell clones with a surface phenotype similar to that of CLL^{14,15}. These clones are detectable at low cell numbers in otherwise healthy individuals using sensitive 6 or 8 color flow cytometry analysis. The condition was recognized when this technology revealed clonal B-cell expansions in adults, particularly, in a series of environmental health studies¹⁴ and in studies of asymptomatic members of CLL kindred¹⁶. MBL can be stratified by the absolute B-cell count into “low count” (normal ALC) and “clinical” (lymphocytosis present) MBL. Individuals with low count MBL have a very low risk of progression. However, for those with lymphocytosis, the risk of progression to CLL is approximately 1–2% per year¹⁷. The majority of patients identified from the general population have low count MBL. We and others have reported MBL in 13–18% of first-degree relatives of CLL patients in high-risk families (i.e., families with 2 or more members with CLL)^{16,18}. This can be compared to 3–5% in the general population using comparable laboratory detection methods and suggests that MBL is a marker of inherited predisposition to CLL^{19–21}. Goldin *et al.* recently reported results of multi-parameter flow cytometry in 505 first-degree relatives with no personal history of lymphoproliferative disease from 140 families having at least two cases of CLL²². Consistent with earlier data, 17% of relatives had MBL. Age was the most important determinant in which the probability for developing MBL by age 90 is estimated to be as high as 61%. MBL clustered in certain families; however, clustering was independent of the number of known CLL cases in a family with two or more CLL members. As is the case

with CLL, males had a significantly higher risk for MBL (19.5%) than did females (15.2%, $p=0.04$). The high rate of MBL and male predominance in high-risk CLL families implies a shared inherited risk with CLL. Most of the MBL cases in high-risk CLL families have low-count MBL. Further immunophenotypic, cytogenetic, and gene expression studies of MBL clones from these families show heterogeneity of the clones although the majority has features of good prognosis of CLL, such as a high prevalence of 13q deletions²³. If MBL is an early step in the process of development of CLL, then germline genes are likely to be acting early in leukemogenesis with more oncogenic events required before CLL develops.

Linkage Studies of CLL

Partly based on the strong evidence for a high familial contribution to CLL, investigators have performed family-based linkage studies of kindred that have two or more related individuals with CLL. These studies represent unbiased genome screens, which can assess the correlation between germline genetic markers and the inheritance of CLL. They are unbiased with regard to the fact that prior biological knowledge is not required for genomic localization. In addition, they are amenable to identifying moderately and rare variants that have large effect sizes²⁴. To date, three linkage studies have been performed^{25–27}. The two earlier linkage studies^{25,26} identified interesting regions but none achieved statistical significance. The more recent study²⁷ and the largest of the three included 206 CLL families with 7,495 genotyped SNPs. Significant linkage was identified on chromosomal band 2q21.2. Other regions of interest from this study were identified on chromosomal bands 5q23.2, 6p22.1, 11q12.1, and 18q21.1. The 2q21.2 region contains the chemokine receptor (*CXCR4*) gene, which plays a key role in B lymphopoiesis, and the region on 6q22.1 corresponds to the major histocompatibility locus. Several factors now favor population studies over additional follow-up linkage studies in CLL families to further evaluate and validate these findings. These include the rarity of CLL, difficulty in collecting CLL families, as well as the theoretical findings by Risch and Merikangas²⁸ that studies of unrelated individuals with over a million variants genotyped could be more powerful than studies of related individuals with only 10,000 variants genotyped. However, in a follow-up sequencing study of the top linkage signal, the coding and splice site regions within *CXCR4* were sequenced in 188 familial CLL patients and 213 controls²⁹. There was no evidence of modestly frequent variants with large effect size, but instead rare functional coding variants with frequencies of 0.5% in the cases but not in the controls. These findings suggest rare variants can contribute to CLL risk.

Candidate gene studies

Candidate gene studies are hypothesis driven studies based on partial or associative relationships between genes and possible involvement in a related pathway underlying a disease, such as CLL. Prior to 2007, a large number of candidate gene studies were performed in CLL (reviewed elsewhere³⁰), but with little durable success. Most often, they were typically underpowered, included no independent validation, were too liberal in statistical significance, and had the fundamental limitation that they failed to include heretofore unknown genes. Further, given the available technology at the time, the evaluation of the selected genes was not comprehensive but rather based on a small number of putative functional variants from each gene. As a result, most of the initial findings had not been replicated in larger and better powered studies. A more recent larger candidate gene study was performed in which 692 CLL cases and 748 controls of European descent from Spain had 613 single nucleotide polymorphisms (SNPs) successfully genotyped from 172 genes involved in cancer biology (e.g., DNA repair, apoptosis)³¹. Significant findings ($P<10^{-5}$) were observed for 4 SNPs in *CCNH*, *APAF1*, *IL16*, and *CASP8*, two of which *APAF1* and *CASP8* are both involved in apoptosis. Interestingly, SNPs from *CASP8* and

CCNH were also evaluated in an earlier study with 992 CLL cases and 2,707 controls of European descent from the United Kingdom, but were found not to be statistically significant in the prior study³². The discrepancy suggests possible CLL heterogeneity across the two studies or either set of results are due to chance. This latter point underscores the need for improved gene coverage, larger samples sizes, and consortial participation characteristic of more recent studies that are systematically resulting in a more accurate list of susceptibility genes.

Genome-wide Association Studies (GWAS)

With the results of the HapMap project and the development of massively parallel array-based genotyping technologies, genome-wide association studies (GWAS) became feasible in 2007 nearly 10 years from the original hypothesis of Risch and Merikangas²⁸. Similar to linkage studies, these are unbiased genome screens to identify common risk variants with modest effect sizes. The practice of GWAS has resulted in strong rigorous criteria for stringent statistical significance and incorporation of validation samples to declare a novel susceptibility locus. As a result, GWAS discovered loci have demonstrated more consistent replication in independent studies. The first GWAS of CLL was performed in 2008 and included 505 CLL cases and 1,438 controls with European ancestry from the United Kingdom who had 299,983 SNPs passing genotyping quality control³³. Six loci were identified and validated on chromosomal bands 2q13, 2q37.1, 6p25, 11q24, 15q23, and 19q13. Additional independent studies validated 5 of the 6 loci with the 19q13 locus having inconsistent findings^{34,35}. A follow-up analysis of the initial GWAS identified 5 additional susceptibility loci on 2q37.3, 8q24.21, 15q25.2, 15q21.3, and 16q24.1 for a total of 11 novel CLL susceptibility loci^{36,37}. The second GWAS of CLL was conducted in a study of individuals of European ancestry from the San Francisco Bay area and included 148 CLL cases and 592 controls which had 339,528 SNPs successfully genotyped³⁸. While no novel CLL loci were identified in this GWAS, the 6p23 and 11q24 regions identified by the first GWAS were validated. The third GWAS of CLL, led by Slager *et al.*, included 407 CLL cases of which 102 were familial CLL cases, and 296 controls, all of whom were of European descent and had 827,777 successfully genotyped SNPs³⁹. Independent variants located in the *IRF8* gene were identified and validated within the 16q24.1 locus. Further, this third GWAS study performed a subset analysis of familial CLL cases versus controls and identified and validated a novel susceptibility locus within the 6p21.3 region, which includes the major histocompatibility locus; this locus had been also identified in a prior linkage study²⁷. These results possibly suggest a stronger effect of the genetic variants in familial CLL cases compared to that in sporadic CLL cases.

For CLL (and may other diseases), the number of discovered GWAS loci increases with larger sample sizes⁴⁰. The initial three CLL GWASs included between 148 and 505 CLL cases. A meta-analysis that combined the data from these three studies for a total of 1,121 CLL cases and 3,745 controls of European descent identified and validated one additional novel region on 6p21.31 that achieved genome-wide significance⁴¹. This finding was subsequently replicated in an independent sample of 1,196 CLL cases and 1,699 controls⁴².

By far the largest GWAS of CLL to date, was recently undertaken that included a 3-stage study of 3,100 CLL cases and 7,667 controls⁴³ from 26 international studies. From this endeavor, ten independent SNPs for CLL risk in nine novel loci were discovered on chromosomal bands 10q23.31, 18q21.33, 11p15.5, 4q25, 2q33.1, 9p21.3, 18q21.32, 15q15.1, and 2p22.2, as well as one independent SNP at an established locus on 2q13. A number of these novel loci are located in or near genes involved in apoptosis. Of the prior 13 GWAS discovered loci, all were replicated except one on chromosomal band 15q25.2 in the stage 1 of the study, which had 2,179 CLL cases and 6,221 controls genotyped; it was

marginally significant in the meta-analysis ($P=0.03$). This could be a real finding with modest effect size that exhibits heterogeneity. Further studies are needed to elucidate this finding. Including the 15q25.2 locus, we now have a total of 25 independent SNPs from 22 novel loci (Table 1). Interestingly, only one of the prior candidate gene study findings of *CASP8* replicated in the GWAS study. Taken together, these 25 GWAS-discovered variants can explain approximately ~17% of the genetic heritability of incident CLL.

Post-GWAS studies

The CLL GWAS discovered SNPs do not directly account for the underlying biological effect of the susceptibility allele, but instead surrogates for one or more underlying susceptibility haplotypes. Interestingly, many of the GWAS SNPs are located in intergenic regions. A more comprehensive evaluation of the regions and better understanding of regulatory mechanisms is needed to identify potentially functional genetic variants for laboratory evaluation. Strategies to unravel the underlying biology of these regions include fine-mapping studies, expression quantitative loci (eQTL) studies, and multi-ethnic studies, as well as studies assessing GWAS discovered variants in MBL, the precursor to CLL.

Fine-mapping studies survey the often large number of variants genotyped within a small genomic region of interest in order to evaluate the linkage disequilibrium pattern, particularly between the GWAS marker SNPs and highly correlated surrogates that could be functionally important. Two fine-mapping studies of CLL GWAS discovered regions have been published so far: one for the *IRF4* gene located on chromosome 6p25.3⁴⁴ and one for the *IRF8* gene located on chromosome 16q24.1⁴⁵. In both of these studies, the GWAS discovered variant was located within non-coding regions of the genes (the 3'-untranslated region [3' UTR] of *IRF4* and intron 3 of *IRF8*, respectively). Comprehensive sequencing of each region was performed, and subsequent work identified additional variants associated with CLL risk in the 3' untranslated region (3' UTR) of each gene, which are generally not translated and alter gene expression. Future biological studies can now use the ENCODE⁴⁶ data set along with other data to generate detailed hypotheses for further laboratory studies needed to understand the underlying biological effects of one or more susceptibility alleles.

To determine if CLL GWAS discovered variants alter gene expression, expression quantitative trait linkage (eQTL) analysis has emerged as useful tool, especially in CLL^{33,39,44,45,47}. A number of promising correlations have been identified. However, all of these studies used publically available data from unselected samples. Future studies correlating gene expression profiles in CLL cells with germline variant data are needed to determine the specificity of the findings. Epigenetic studies may also help to identify the molecular basis of susceptibility variants that are located outside of the gene coding regions.

Multi-ethnic studies provide insight into sharing of alleles across populations. Although the incidence of CLL varies substantially across ethnicities, with the highest rates in European populations and the lowest rates in Asians, the clinical and pathological features of CLL have been shown to be similar^{48,49}. Two studies have assessed GWAS discovered SNPs in Chinese⁴⁸ and African American⁵⁰ ethnicities. In the Chinese study, 71 CLL cases and 1,273 controls had 6 GWAS SNPs genotyped; of which 3 resulted in significant associations with CLL risk (rs872071, rs13397985, rs17483466). In the African American study, 110 CLL African American cases were genotyped for 15 GWAS discovered SNPs. The CLL allele frequencies from these SNPs were then compared to those publically available African American control samples (HapMap and published African American GWAS of prostate cancer). None of the SNPs showed significant associations ($P>0.05$) with CLL risk in African Americans. It is well-known that allele frequencies vary across ethnicities, and this is exemplified in the findings from these two studies. Table 2 lists 4 CLL GWAS SNPs in

which allele frequencies are available across the different ethnic control groups. In all cases, the allele frequencies of the risk alleles in the non-European studies are lower than that in the European studies. Interestingly, migration studies have shown that Asian subjects, e.g., retain the lower CLL rates characteristic of their country of origin when they migrate to the West^{51–53}. This is consistent with a genetic contribution to CLL etiology and lower frequencies of risk alleles observed in Asians would be consistent with the lower CLL rates, although larger studies are needed to verify these early observations. Moreover, new CLL GWAS are needed in different populations to identify novel loci specific to those ethnic groups surveyed as well as identify SNPs that are common susceptibility loci across populations.

Two studies have evaluated the variants discovered in CLL GWAS in studies designed to evaluate MBL^{39,54}, the precursor condition to CLL. The larger study⁵⁴ had assessed 10 GWAS variants in a meta-analysis of 419 MBL cases and 1,753 controls obtained from three studies. Significant associations ($P < 0.05$) were observed in 6 of the 10 evaluated variants. Three of the non-significant variants had elevated effects suggesting possible statistical power issue due to small sample size. Only the rs11083846 on 19q13.3 had little evidence of an association. These results from these 9 of 10 SNPs further support that MBL is a precursor to CLL. However, because not all individuals with MBL progress to CLL, further studies are needed to determine perhaps additional genetic and non-genetic factors associated with progression to CLL.

Clinical direction

The clinical utility of GWAS-discovered variants has yet to be determined in rigorous studies. As described earlier, first-degree relatives of CLL patients have an 8.5-fold relative risk for developing CLL and are also at increased risk for developing other indolent forms of NHL compared to relatives of matched controls¹³. With the exception of the variants in the major histocompatibility locus on 6p21.3, the GWAS discovered variants have been identified in an unselected sample of CLL cases (i.e., regardless of family history of CLL). Risk-prediction models that incorporate the GWAS variants along with ethnicity, family history, and sex may identify those with greater risk. Studies are planned or underway to determine if the now 25 GWAS-discovered SNPs stratify CLL cases based on clinical outcome. Exome sequencing studies to identify the somatic mutations in CLL tumor cells will complement these efforts⁵⁵. Finally, with the detection of a catalog of germline susceptibility alleles for CLL risk, there is an opportunity to assess whether these variants are also pointing towards possible drug targets for prevention or therapeutic studies.

Summary and Future Directions

The 22 GWAS-discovered loci provide new leads to focus our efforts in understanding genetic susceptibility of CLL. These new findings focus our attention on common alleles, ones that may have smaller effect sizes. Several of the known loci have candidate genes with clear roles in B-cell biology or the apoptosis pathway. Although additional individual variants may be discovered with larger studies, a more integrated approach is needed to investigate the underlying biology of CLL susceptibility. This may include pathway studies and integration of other data types (e.g., gene expression, epigenetics, or protein expression). Identification of rare variants through next generation sequencing will also be another research avenue; however, this will require very large studies or a circle back to pedigree studies. Studies in populations other than those of European ancestry also provide an opportunity to not only discover novel loci but also hone in on the loci that are shared across populations, but perhaps varying in frequency of the susceptibility alleles. Similarly, as larger studies are being conducted on closely related hematological malignancies (e.g., other

non-Hodgkin lymphoma subtypes), shared SNPs, loci, or pathways may be identified that potentially may shed light on shared familiarity of these conditions. Finally, none of the aforementioned studies replace the need for *in vitro* or *in vivo* functional studies. The GWAS discovered findings provide a logical first step for functional studies, which should point us towards a more comprehensive understanding of the genetic contributions to CLL.

Acknowledgments

Supported in part by the National Cancer Institute/NIH Grants R01 CA11844, R01 CA92153; the Intramural research program of the NIH, NCI.

References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin.* 2013; 63:11–30. [PubMed: 23335087]
2. Hallek M, Cheson BD, Catovsky D, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood.* 2008; 111:5446–56. [PubMed: 18216293]
3. Abrisqueta P, Pereira A, Rozman C, et al. Improving survival in patients with chronic lymphocytic leukemia (1980–2008): the Hospital Clinic of Barcelona experience. *Blood.* 2009; 114:2044–50. [PubMed: 19553638]
4. Brenner H, Gondos A, Pulte D. Trends in long-term survival of patients with chronic lymphocytic leukemia from the 1980s to the early 21st century. *Blood.* 2008; 111:4916–21. [PubMed: 18309034]
5. Call TG, Phylilly RL, Noel P, et al. Incidence of chronic lymphocytic leukemia in Olmsted County, Minnesota, 1935 through 1989, with emphasis on changes in initial stage at diagnosis. *Mayo Clinic Proc.* 1994; 69:323–8.
6. Howlader, N.; Noone, AM.; Krapcho, M., et al. National Cancer Statistics Review. National Cancer Institute; Bethesda: 2012. SEER Cancer Statistics Review, 1975–2010.
7. Linet, MS.; Devesa, SS.; Morgan, GJ. The Leukemias. In: Schottenfeld, D.; Fraumeni, JF., Jr, editors. *Cancer Epidemiology and Prevention.* Oxford University Press, Inc; 2006.
8. Albright F, Teerlink C, Werner TL, Cannon-Albright LA. Significant evidence for a heritable contribution to cancer predisposition: a review of cancer familiarity by site. *BMC Cancer.* 2012; 12:138. [PubMed: 22471249]
9. Yuille MR, Matutes E, Marossy A, et al. Familial chronic lymphocytic leukaemia: a survey and review of published studies. *Br J Haematol.* 2000; 109:794–9. [PubMed: 10929032]
10. Sgambati, M.; Linet, MS.; Devesa, SS. Chronic Lymphocytic Leukemia: Epidemiological, Familial, and Genetic Aspects. In: Cheson, BD., editor. *Chronic Lymphoid Leukemias.* Marcel Dekker, Inc; 2001. p. 33-62.
11. Tjonnfjord GE, Jonsson V, Ly BE, Johannesen TB. Familial occurrence of chronic lymphocytic leukaemia in Norway. *Tidsskr Nor Laegeforen.* 2012; 132:2060–3. [PubMed: 23038196]
12. 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; 104:1850–4. [PubMed: 15161669]
13. Goldin LR, Bjorkholm M, Kristinsson SY, Turesson I, Landgren O. Elevated risk of chronic lymphocytic leukemia and other indolent non-Hodgkin's lymphomas among relatives of patients with chronic lymphocytic leukemia. *Haematologica.* 2009; 94:647–53. [PubMed: 19286886]
14. Marti G, Abbasi F, Raveche E, et al. Overview of monoclonal B-cell lymphocytosis. *Br J Haematol.* 2007; 139:701–8. [PubMed: 18021084]
15. Marti GE, Rawstron AC, Ghia P, et al. Diagnostic criteria for monoclonal B-cell lymphocytosis. *Br J Haematol.* 2005; 130:325–32. [PubMed: 16042682]
16. Marti GE, Carter P, Abbasi F, et al. B-cell monoclonal lymphocytosis and B-cell abnormalities in the setting of familial B-cell chronic lymphocytic leukemia. *Cytometry B Clin Cytom.* 2003; 52:1–12. [PubMed: 12599176]

17. Shanafelt TD, Kay NE, Rabe KG, et al. Brief report: natural history of individuals with clinically recognized monoclonal B-cell lymphocytosis compared with patients with Rai 0 chronic lymphocytic leukemia. *J Clin Oncol*. 2009; 27:3959–63. [PubMed: 19620484]
18. Rawstron AC, Yuille MR, Fuller J, et al. Inherited predisposition to CLL is detectable as subclinical monoclonal B-lymphocyte expansion. *Blood*. 2002; 100:2289–90. [PubMed: 12239136]
19. Dagklis A, Fazi C, Sala C, et al. The immunoglobulin gene repertoire of low-count chronic lymphocytic leukemia (CLL)-like monoclonal B lymphocytosis is different from CLL: diagnostic implications for clinical monitoring. *Blood*. 2009; 114:26–32. [PubMed: 19029437]
20. Ghia P, Prato G, Scielzo C, et al. Monoclonal CD5+ and CD5- B-lymphocyte expansions are frequent in the peripheral blood of the elderly. *Blood*. 2004; 103:2337–42. [PubMed: 14630808]
21. Rawstron AC, Green MJ, Kuzmicki A, et al. Monoclonal B lymphocytes with the characteristics of “indolent” chronic lymphocytic leukemia are present in 3.5% of adults with normal blood counts. *Blood*. 2002; 100:635–9. [PubMed: 12091358]
22. Goldin LR, Lanasa MC, Slager SL, et al. Common occurrence of monoclonal B-cell lymphocytosis among members of high-risk CLL families. *Br J Haematol*. 2010; 151:152–8. [PubMed: 20738309]
23. Lanasa MC, Allgood SD, Slager SL, et al. Immunophenotypic and gene expression analysis of monoclonal B-cell lymphocytosis shows biologic characteristics associated with good prognosis CLL. *Leukemia*. 2011; 25:1459–66. [PubMed: 21617698]
24. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008; 9:356–69. [PubMed: 18398418]
25. Goldin LR, Ishibe N, Sgambati M, et al. A genome scan of 18 families with chronic lymphocytic leukaemia. *Br J Haematol*. 2003; 121:866–73. [PubMed: 12786797]
26. Sellick GS, Webb EL, Allinson R, et al. A high-density SNP genomewide linkage scan for chronic lymphocytic leukemia-susceptibility loci. *Am J Hum Genet*. 2005; 77:420–9. [PubMed: 16080117]
27. Sellick GS, Goldin LR, Wild RW, et al. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. *Blood*. 2007; 110:3326–33. [PubMed: 17687107]
28. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science*. 1996; 273:1516–7. [PubMed: 8801636]
29. Crowther-Swanepoel D, Qureshi M, Dyer MJ, et al. Genetic variation in CXCR4 and risk of chronic lymphocytic leukemia. *Blood*. 2009; 114:4843–6. [PubMed: 19812382]
30. Slager SL, Kay NE, Fredericksen ZS, et al. Susceptibility genes and B-chronic lymphocytic leukaemia. *Br J Haematol*. 2007; 139:762–71. [PubMed: 18021089]
31. Enjuanes A, Benavente Y, Bosch F, et al. Genetic variants in apoptosis and immunoregulation-related genes are associated with risk of chronic lymphocytic leukemia. *Cancer Res*. 2008; 68:10178–86. [PubMed: 19074885]
32. Rudd MF, Sellick GS, Webb EL, Catovsky D, Houlston RS. Variants in the ATM-BRCA2-CHEK2 axis predispose to chronic lymphocytic leukemia. *Blood*. 2006; 108:638–44. [PubMed: 16574953]
33. 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; 40:1204–10. [PubMed: 18758461]
34. Slager SL, Goldin LR, Strom SS, et al. Genetic susceptibility variants for chronic lymphocytic leukemia. *Cancer Epidemiol Biomarkers Prev*. 2010; 19:1098–102. [PubMed: 20332261]
35. Crowther-Swanepoel D, Mansouri M, Enjuanes A, et al. Verification that common variation at 2q37.1, 6p25.3, 11q24.1, 15q23, and 19q13.32 influences chronic lymphocytic leukaemia risk. *Br J Haematol*. 2010; 150:473–9. [PubMed: 20553269]
36. 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; 42:132–6. [PubMed: 20062064]

37. 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; 154:229–33. [PubMed: 21554262]
38. 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; 42:661–4. [PubMed: 20639881]
39. 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; 117:1911–6. [PubMed: 21131588]
40. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet.* 2012; 90:7–24. [PubMed: 22243964]
41. Slager SL, Skibola CF, Di Bernardo MC, et al. Common variation at 6p21.31 (BAK1) influences the risk of chronic lymphocytic leukemia. *Blood.* 2012; 120:843–6. [PubMed: 22700719]
42. Slager SL, Camp NJ, Conde L, et al. Common variants within 6p21.31 locus are associated with chronic lymphocytic leukaemia and, potentially, other non-Hodgkin lymphoma subtypes. *Br J Haematol.* 2012; 159:572–6. [PubMed: 23025533]
43. Berndt SI, Skibola CF, Joseph V, et al. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nat Genet.* 2013
44. Crowther-Swanepoel D, Broderick P, Ma Y, et al. Fine-scale mapping of the 6p25.3 chronic lymphocytic leukaemia susceptibility locus. *Hum Mol Genet.* 2010; 19:1840–5. [PubMed: 20123861]
45. Slager SL, Achenbach SJ, Asmann YW, et al. Mapping of the IRF8 gene identifies a 3 UTR variant associated with risk of chronic lymphocytic leukemia but not other common non-Hodgkin lymphoma subtypes. *Cancer Epidemiol Biomarkers Prev.* 2013; 22:461–6. [PubMed: 23307532]
46. Maher B. ENCODE: The Human Encyclopaedia. *Nature.* 2012; 489:57–74. [PubMed: 22955616]
47. Sille FC, Thomas R, Smith MT, Conde L, Skibola CF. Post-GWAS functional characterization of susceptibility variants for chronic lymphocytic leukemia. *PLoS One.* 2012; 7:e29632. [PubMed: 22235315]
48. Lan Q, Au WY, Chanock S, et al. Genetic susceptibility for chronic lymphocytic leukemia among Chinese in Hong Kong. *Eur J Haematol.* 2010; 85:492–5. [PubMed: 20731705]
49. Coombs CC, Falchi L, Weinberg JB, Ferrajoli A, Lanasa MC. Chronic lymphocytic leukemia in African Americans. *Leuk Lymphoma.* 2012; 53:2326–9. [PubMed: 22646816]
50. Coombs CC, Rassenti LZ, Falchi L, et al. Single nucleotide polymorphisms and inherited risk of chronic lymphocytic leukemia among African Americans. *Blood.* 2012; 120:1687–90. [PubMed: 22745306]
51. Parkin DM. International variation. *Oncogene.* 2004; 23:6329–6340. [PubMed: 15322508]
52. Caligaris-Cappio F. What are we learning from familial chronic lymphocytic leukemia? *Leukemia Res.* 2002; 26:779–780. [PubMed: 12127549]
53. Pang JWY, Cook LS, Schwartz SM, Weiss NS. Incidence of Leukemia in Asian Migrants to the United States and Their Descendants. *Cancer Causes Control.* 2002; 13:791–795. [PubMed: 12462543]
54. Crowther-Swanepoel D, Corre T, Lloyd A, et al. Inherited genetic susceptibility to monoclonal B-cell lymphocytosis. *Blood.* 2010; 116:5957–60. [PubMed: 20855867]
55. Gunnarsson R, Mansouri L, Rosenquist R. Exploring the genetic landscape in chronic lymphocytic leukemia using high-resolution technologies. *Leuk Lymphoma.* 2013; 54:1583–90. [PubMed: 23167608]

Table 1

GWAS discovered loci and SNPs

Locus	Nearest gene	Published SNP	Location	Refs
2p22.2	<i>QPCT, PRKD3</i>	rs3770745	37,449,593	43
2q13	<i>ACOXL</i>	rs17483466 rs13401811	111,513,929 111,332,575	33,43
2q33.1	<i>CASP10, CASP8</i>	rs3769825	201,819,625	43
2q37.1	<i>SP140</i>	rs13397985	230,799,467	33
2q37.3	<i>FARP2</i>	rs757978	242,019,774	36
4q25	<i>LEF1</i>	rs898518	109,236,273	43
6p25.3	<i>IRF4</i>	rs872071	356,064	33
6p21.32	<i>HLA</i>	rs674313	32,686,060	39
6p21.31	<i>BAK1</i>	rs210142	33,654,815	41
8q24.21	<i>MYC</i>	rs2456449	128,262,163	36
9p21.3	<i>CDKN2B-AS1</i>	rs1679013	22,196,987	43
10q23.31	<i>ACTA2, FAS</i>	rs4406737	90,749,704	43
11p15.5	<i>C11orf21, TSPAN32</i>	rs7944004	2,267,728	43
11q24.1	<i>SCN3B</i>	rs735665	122,866,607	33
15q15.1	<i>BMF</i>	rs8024033	38,190,949	43
15q21.3	<i>RFXDC2</i>	rs7169431	54,128,188	36
15q23	<i>RPLP1</i>	rs7176508	67,806,044	33
15q25.2	<i>CPEB1</i>	rs783540	81,051,763	37
16q24.1	<i>IRF8</i> <i>IRF8</i>	rs2292982 rs305061	84,502,324 84,533,160	36,39
18q21.32	<i>PMAIP1</i>	rs4368253	55,773,267	43
18q21.33	<i>BCL2</i>	rs4987855 rs4987852	58,944,529 58,944,901	43
19q13.3	<i>PRKD2</i>	rs11083846	51,899,494	33

Table 2

Risk allele frequencies of GWAS discovered loci in controls from three ethnicities

Locus	Nearest gene	Published SNP	risk allele	Caucasian ^a	Chinese ^b	African American ^c
2q13	ACOXL	rs17483466	G	0.200	0.013	0.060
2q27.1	SP140	rs13397985	G	0.180	0.003	0.100
6p25.3	IRF4	rs872071	G	0.510	0.310	0.090
11q24.1	SCN3B	rs735665	A	0.190	0.005	0.050

^a ref 43;

^b ref 47;

^c ref 49