

## Chronic lymphocytic leukemia - Section 5

# Genetic predisposition to chronic lymphocytic leukemia

**Philip J. Law, Richard S. Houlston**

*Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, Surrey, United Kingdom*

### Take home messages

- Chronic lymphocytic leukemia (CLL) is characterized by having one of the strongest familial risks of any cancer.
- Genome-wide association studies have identified common variants mapping to over 40 regions of genome that influence the risk of developing sporadic CLL. Sequencing of familial CLL has implicated rare loss-of-function mutations in shelterin complex genes in CLL predisposition.
- As well as providing new insights in the developmental basis of CLL, the cancer gene discovery initiatives have potential to inform the development of new therapeutic agents.

### Introduction

Chronic lymphocytic leukemia (CLL) is an indolent B-cell malignancy that has a strong inherited component, as evidenced by the 8-fold increased risk seen in relatives of CLL patients.<sup>\*1</sup> Until recently, inherited genetic basis to CLL was unknown. Our understanding of CLL genetics has been transformed by the genome-wide association studies (GWAS) of CLL performed over the last 10 years.<sup>\*2,3,\*4,5</sup> These GWAS have provided the first direct evidence for inherited susceptibility to CLL identifying common variants at over 40 independent genomic regions influencing risk of sporadic disease. In addition to common genetic variation influencing risk, high-throughput sequencing studies of CLL families have established a key role for rare disruptive mutations as determinants of disease susceptibility. Besides providing evidence for genetic susceptibility to CLL, the genetic regions and genes risk identified by these analyses have provided fresh insights into the biological basis of CLL development.

### Current state of the art

GWAS have so far identified single nucleotide polymorphisms (SNPs) at 43 independent genetic regions that influence the risk of

developing sporadic CLL.<sup>\*2,3,\*4,5</sup> While the risk of CLL associated with each of the GWAS risk SNPs is modest, in concert they have the potential to have more profound effects on an individual's risk of developing CLL. Thus far the currently identified risk SNPs account for 25% of the heritable risk. By fitting all SNPs from GWAS simultaneously using statistical modeling has shown that the estimated heritability of CLL attributable to all common variation is 34%, thus having potential to explain 57% of the overall familial risk<sup>\*4</sup> and confirming the long held belief that a significant part of the heritable risk of CLL is polygenic in nature.<sup>6</sup>

To the extent that they have been deciphered, most cancer GWAS risk regions map to noncoding regions of the genome and influence disease risk by altering gene regulation.<sup>\*7</sup> This is also the case for CLL with over 75% of risk regions showing enrichment of active promoters and/or enhancers when assessed by H3K27ac, H3K4me3, and H3K4me1 ChIP-seq marks. Moreover, CLL risk regions are enriched for CLL-related regulatory elements which are CLL-specific or show differential regulation across CLL and B-cell development.<sup>\*4</sup> Quantitative trait locus analysis in conjunction with data from H3K27ac, ATAC-seq, and DNA methylation profiling are consistent with risk regions mediating their effects by influencing chromatin activity. The identified loci show an over-representation of transcription factor (TF) binding. Several of the TFs mapping to risk regions have well-established roles in B-cell function; for example, OCT2, IRF4, and RUNX3 being targeted for hypomethylation in B-cells. MYC is a direct target of IRF4 in activated B cells, with IRF4 itself being a direct target of MYC transactivation. In this respect, it is noteworthy that genetic variants at IRF4 and MYC are recognized factors for CLL pathogenesis. Collectively, findings from these analyses are consistent with CLL GWAS risk SNPs mapping within regions of active chromatin state that exert effects on B-cell cis-regulatory networks. Investigating the genetic pathways between the gene products in proximity to the GWAS SNPs, it has been shown that gene products are primarily involved in immune response, B-cell receptor (BCR)-mediated signaling, apoptosis, and maintenance of chromosome integrity, as well as interconnectivity between the

*Funding/support:* None.

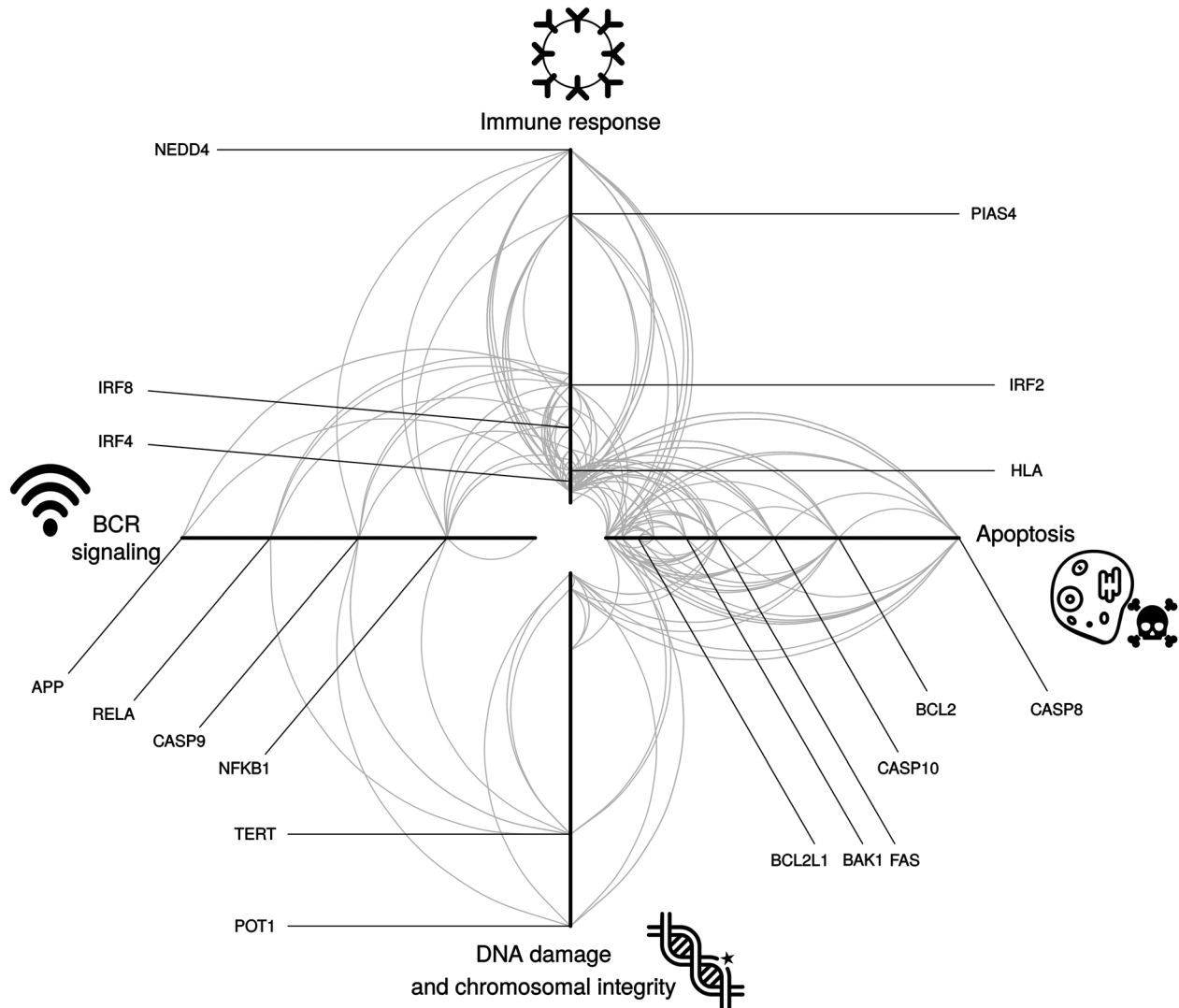
*Disclosure:* The authors have indicated they have no potential conflicts of interest to disclose.

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the European Hematology Association. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*HemaSphere* (2019) 3:S2

Received: 28 January 2019 / Accepted: 23 February 2019

**Citation:** Law PJ, Houlston RS. Genetic Predisposition to Chronic Lymphocytic Leukemia. *HemaSphere*, 2019;3:S2. <http://dx.doi.org/10.1097/HS9.000000000000194>



**Figure 1. Genes implicated by genome-wide association studies as having a role in defining the development of CLL.** Each arm corresponds to a defined biological process. Each arc represents an interaction between 2 proteins, and the distance from the center of the plot corresponds to a greater number of protein-protein interactions (higher degree of the node). Selected proteins known to be involved in CLL risk are indicated. CLL = chronic lymphocytic leukemia.

gene products all being central to B-cell development (Fig. 1). It is notable that Ibrutinib (a BTK inhibitor) and Idelalisib (a PI3KCD inhibitor) mediate their effects through interference of BCR signaling, and Venetoclax targets the antiapoptotic behavior of BCL-2.

Families segregating CLL have provided evidence for Mendelian susceptibility; however, until recently, the identification of rare alleles with large effects has been elusive. The identification of this class of susceptibility is especially important because mutations are causal and provide direct insight to cancer biology, in contrast to GWAS associations. By performing whole-exome sequencing of CLL families, loss-of-function mutations in Protection of Telomeres 1 (POT1) and other components of the shelterin complex have been demonstrated.<sup>8,9</sup> As well as providing support for the role of rare variants these findings further highlight telomere dysregulation as a key process in CLL development. Moreover, they extend the spectrum of cancer associated with inherited mutations in these genes. It is, however, likely that shelterin complex gene mutations confer cancer risks analogous to those associated with ATM heterozygosity<sup>10</sup> or CHEK2 for breast cancer.<sup>11</sup> Nevertheless, because the dysregu-

lation of telomere protection has been identified as a target for potential therapeutic intervention in CLL, it may be possible that early identification of mutation carriers will facilitate improvements in future disease management.

### Future perspectives

Recent studies have provided the first direct evidence for inherited predisposition to CLL. As well as providing new insights in the developmental basis of CLL, these gene discovery initiatives have the potential to impact on risk prediction and on the successful development of new therapeutic agents. Since much of the inherited risk of CLL still remains unexplained additional studies based on larger datasets offer an opportunity to identify new risk regions and susceptibility genes. Deciphering the function of GWAS risk loci is an important step toward testable hypotheses regarding the biological processes involved in pathogenesis. Elucidating the mechanisms through which noncoding variants exert their effect is, however, challenging as the genotyped SNP is not generally a strong candidate for causality. While studies to fully elucidate the regulatory mechanisms underpinning risk

regions are in their relative infancy, such endeavors are likely to involve high-throughput systems such as massive parallel reporter assays and exploration of tissue-specific effects in appropriate model systems and CRISPR-Cas9-mediated disruption of candidate regulatory elements.

## References

- \*1. Goldin LR, Pfeiffer RM, Li X, et al. Familial risk of lymphoproliferative tumors in families of patients with chronic lymphocytic leukemia: results from the Swedish Family-Cancer Database. *Blood*. 2004;104:1850–1854.

### Key reference on familial risks of chronic lymphocytic leukaemia

- \*2. Berndt SI, Camp NJ, Skibola CF, et al. Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia. *Nat Commun*. 2016;7:10933.

### First GWAS of chronic lymphocytic leukaemia demonstrating inherited predisposition

3. 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;40:1204–1210.
- \*4. Law PJ, Berndt SI, Speedy HE, et al. Genome-wide association analysis implicates dysregulation of immunity genes in chronic lymphocytic leukaemia. *Nat Commun*. 2017;8:14175.

### Most comprehensive GWAS of chronic lymphocytic leukaemia

5. Speedy HE, Di Bernardo MC, Sava GP, et al. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. *Nat Genet*. 2014;46:56–60.
6. Kleinstern G, Camp NJ, Goldin LR, et al. Association of polygenic risk score with the risk of chronic lymphocytic leukemia and monoclonal B-cell lymphocytosis. *Blood*. 2018;131:2541–2551.
- \*7. Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. *Nat Rev Cancer*. 2017;17:692–704.

### Overview of cancer GWAS

- \*8. Speedy HE, Kinnersley B, Chubb D, et al. Germ line mutations in shelterin complex genes are associated with familial chronic lymphocytic leukemia. *Blood*. 2016;128:2319–2326.

### Report of rare disruptive mutation in shelterin genes as a cause of chronic lymphocytic leukaemia

9. Calvete O, Martinez P, Garcia-Pavia P, et al. A mutation in the POT1 gene is responsible for cardiac angiosarcoma in TP53-negative Li-Fraumeni-like families. *Nat Commun*. 2015;6:8383.
10. Renwick A, Thompson D, Seal S, et al. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet*. 2006;38:873–875.
11. Meijers-Heijboer H, van den Ouweland A, Klijn J, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(\*) 1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet*. 2002;31:55–59.