Leucemia linfática crónica familiar: La predisposición genética de la LLC y la linfocitosis monoclonal de células B

Overview of familial chronic lymphocytic leukaemia: The genetic susceptibility to chronic lymphocytic leukaemia and monoclonal B-cell lymphocytosis

Catovsky D

Divison of Molecular Pathology, The Institute of Cancer Research, London, UK

daniel.catovsky@icr.ac.uk

Palabras claves: LLC familiar, linfocitosis B monoclonal, GWAS.

Keywords: familial CLL, monoclonal B-cell lymphocytosis, GWAS.

Introduction: historical background

Chronic lymphocytic leukaemia (CLL) is the most common form of leukaemia accounting for 30% of all cases in Western countries. Several early epidemiological and cohort studies have documented an increased incidence of the disease and of other lymphoproliferative disorders (LPD) in relatives of patients with CLL (Reviewed in [1] and [2]).

Important early studies by Videbaek in his doctoral thesis[3] and later in a paper[4] and by Gunz[5,6] on familial cases became important landmarks focusing attention on the familial clustering of cases. Familial aggregations of CLL provided the first evidence of an inherited predisposition to the disease.

I personally became interested in the topic in the 1980s. Some years after I treated a patient (1917-1980 in **Figure 1**) I received a short letter from his brother (a physician born in 1925) with the drawing of a family tree (**Figure 1**) asking whether I was aware that three of his uncles had died in New York of CLL. I was not aware of the early literature and in those days we were not taking such detailed family histories as we do today. He suggested that it was an interesting topic for research. A few years later he wrote again and told me that his sister (1919-1994 in **Figure 1**) also developed CLL and subsequently died. He himself was followed at another hospital with a myeloproliferative disorder (MPD). When we received a sample of his blood we could establish that he also had monoclonal B-cell lymphocytosis (MBL).
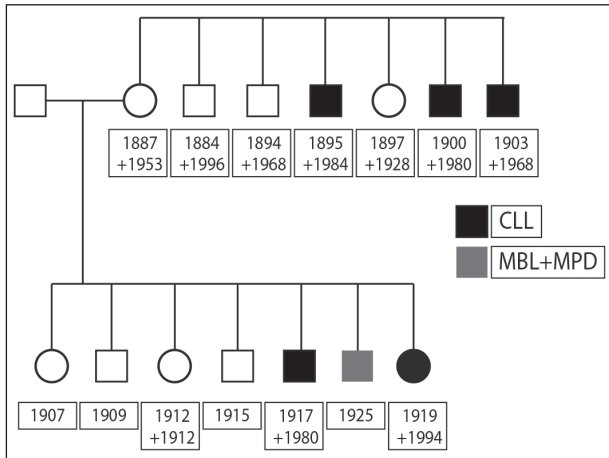
Figure 1. Pedigree of the first CLL family that was brought to my attention in the 1980s

After reading the existing relevant literature[3-6] we decided to start collecting CLL families and established an active international collaboration with colleagues at the USA National Cancer Institute and other countries in order to gain insight into the genetic basis of CLL. We started with a systematic family survey of 268 patients on the basis of a family history questionnaire[1] and in 1999 we formed a Familial CLL Consortium. Examples of some of those family pedigrees are shown in **Figure 2**.
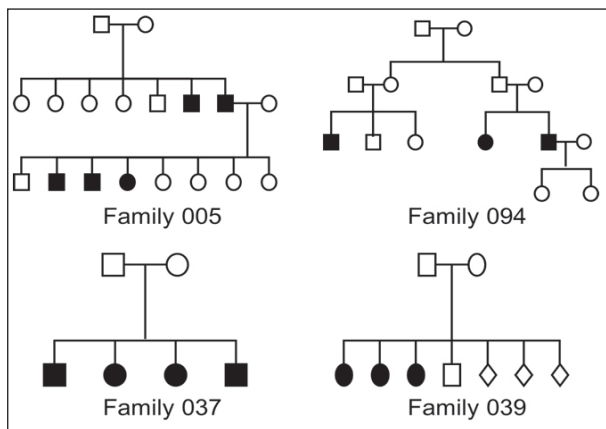

Figure 2. Several of the first pedigrees of our familial CLL study

We also started collecting and storing DNA from blood cells and saliva (germline DNA) in order to be able to carry out later genetic linkage analysis. A few years later we initiated a program of Genome Wide Association Studies (GWAS).

Another important tool for identifying CLL families is the comprehensive medical records and cancer registries available from countries in Northern Europe. For example, from Danish medical records and the Danish Cancer Registry we have been able to follow up the largest CLL family (Pedigree 14) from Videbaek[3] published in 1947. From the original 57 family members Pedigree 14 was extended to 222 family members (**Figure 3**). Overall 10 of them have been diagnosed with CLL, one with T-cell lymphoma and 17 with other cancers, including five with breast cancer. This careful and long follow up provided strong support for an inherited predisposition to CLL, although no research samples were available from Pedigree 14.

Another important study originated from data from the Swedish Family Cancer Database[8]. The familial risk of CLL and other LPD was assessed in 14,336 relatives of 5,918 patients with CLL and in 28,876 first degree relatives of 11,778 normal controls. The first degree relatives of CLL patients (parents, siblings and offspring) had a 7.5-fold increased relative risk (RR) for CLL, a 1.45 RR for non-Hodgkin's lymphoma and a 2.35 risk for Hodgkin's lymphoma[8], suggesting that the heritable component in CLL is shared with other lymphoid disorders. This study is currently being revisited with three times more patients and relatives and a longer follow up. It will also examine the association with other haematological malignancies. The data confirmed that first degree relatives of patients with CLL have a nearly six-fold RR of developing CLL (Dr A.Sud, personal communication).

Comparison between familial and sporadic CLL
In order to investigate whether there were any major differences between sporadic and familial CLL we compared 265 cases of familial CLL with 647 sporadic cases[9].

There were no overt differences in the repertoire and frequency of the variable region of the immunoglobulin heavy chain (*IGHV*) gene usage. Neither was the *IGHV* correlated between affected members of the same family. This pointed to a genetic basis of disease development rather than a shared environmental aetiology.

The only difference observed was in the higher incidence of *IGHV* mutated status in familial cases (68%) than in sporadic ones (47%; p<0.001) and, as expected, there was concordance within the families[9]. No other differences were observed such as age, male-female ratio, incidence of Zap 70+ cells or karyotype abnormalities. We concluded that familial CLL is essentially indistinguishable from sporadic CLL.
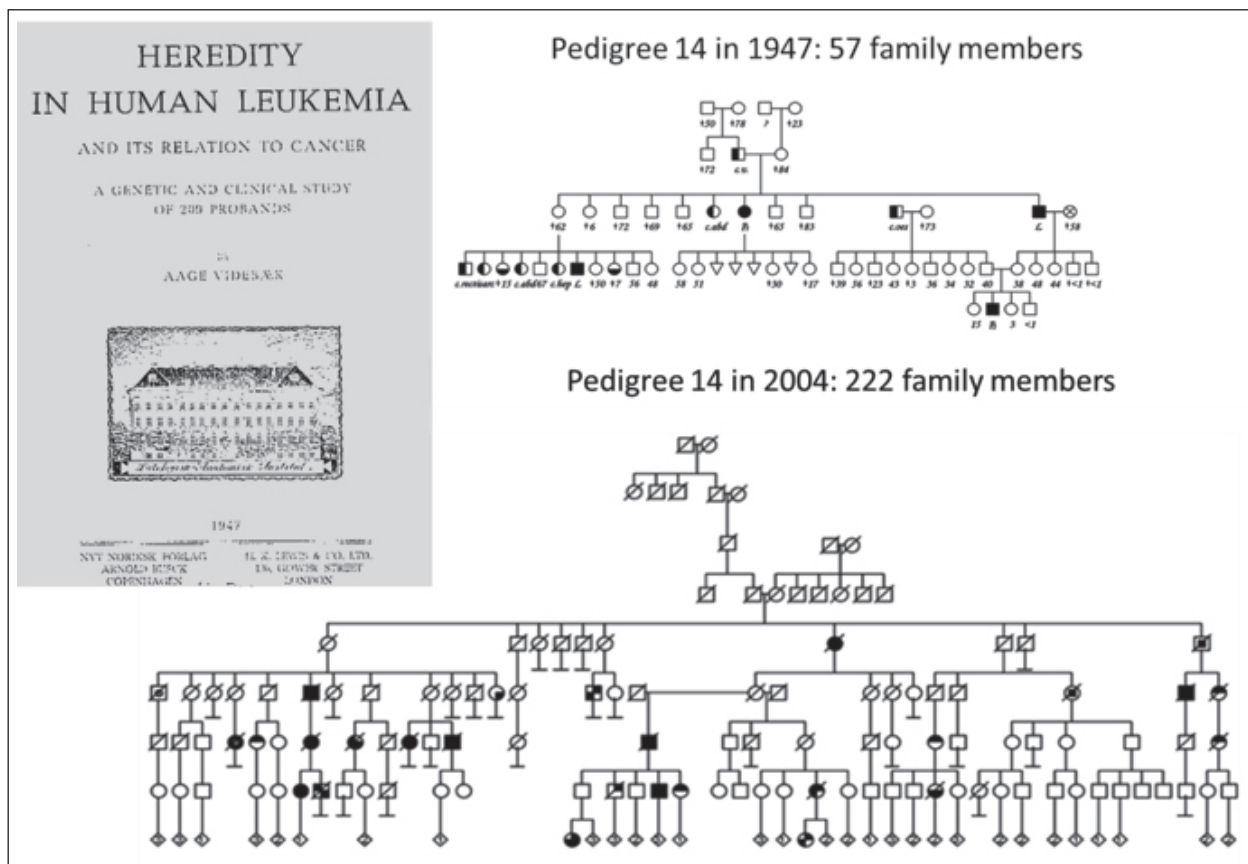
Figure 3. *Pedigree 14 by Videbaek in 1947[3], subsequently expanded in 2004 by Jønsson[7]*

Strategies for identifying CLL predisposition genes

These strategies have evolved over the years in parallel with progress in the technology of DNA analysis, especially regarding single nucleotide polymorphisms (SNPs). Derived (since 2006) from the Human Genome Project and the International HapMap Project, these strategies include: candidate gene studies, linkage analysis of families, genome-wide association studies (GWAS) and, lately, whole-exome sequencing (WES), looking for loss of function mutations in familial cases.

There were many candidate gene studies in CLL including evaluation of more than 100 genes published between 1981 and 2014, but no robust associations with CLL risk were established.

Linkage estimates the genetic distance between loci based on the probability of recombination. The largest study of familial CLL, from our group[10], based on the analysis of 206 pedigrees, failed to identify moderate-high penetrance risk alleles causing CLL. However, linkage in familial cases was found to chromosomes 7q31-q33 and 16q 12.2-q23, a finding which is consistent with the more recent observations of loss-of-function germline mutations in shelterin complex genes[11; see below].

The greatest advances have taken place with GWAS in several other cancers: colorectal, breast, prostate (reviewed in [12]), as well as significant progress in CLL (updated in [13]).

GWAS in CLL

GWAS provides an agnostic and unbiased approach for investigating the genetic basis of complex diseases by examining regions of haplotype sharing in the DNA of affected individuals within each family. The co-segregation of markers with disease in families is greater than expected by chance alone.

GWAS have identified thus far 42 single nucleotide polymorphisms (SNPs) that are low penetrance risk alleles robustly associated with CLL[13]. The majority of these SNPs map to non-coding regions of the

genome and are the most common type of genetic variation among individuals. Each SNP represents a difference in a single DNA building block called a nucleotide. As shown in **Figure 4**, GWAS compare the frequency of genetic variants in CLL cases versus controls. The latest study was based on data from 6,200 cases and 17,598 controls[13]. The common genetic variation thus identified contributes significantly to the heritable risk of CLL following the "common-disease common-variant model" (**Figure 5**).
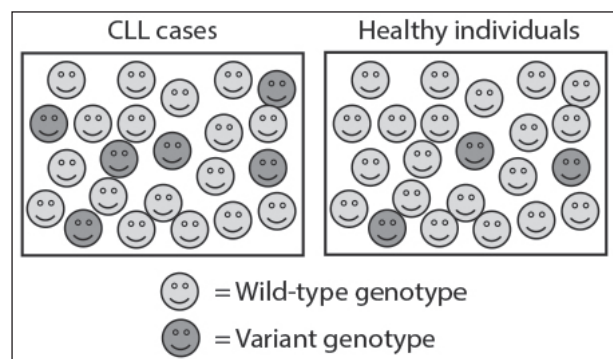

Figure 4. Common risk variants in CLL: genome-wide association studies (GWAS) compare the frequencies of genetic variants in CLL cases versus controls
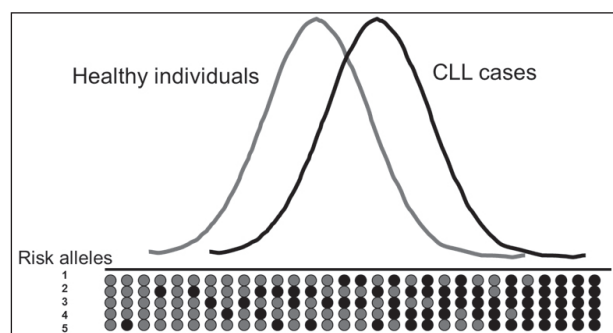

Figure 5. *The common-disease common-variant model for cancer susceptibility with low penetrance genes*

It should be noted that the focus has now moved from families to all cases of CLL as the methodology required large number of samples and strict statistical analysis (differences require a p value of $< 5 \times 10^{-8}$). The inclusion of familial CLL is still important as such cases are enriched in the number of predisposition loci.

The first two GWAS reports identified 10 susceptibility loci for CLL[14,15] and, with international collaborations, these have now been extended to 42 predisposition risk loci[13].

GWAS requires the evaluation c.300,000 tagging SNPs, which are sufficient to capture the majority or all the common genetic variation in the region[12-15]. This large number of samples is required because the SNPs have a high frequency in a given population and confer a low penetrance risk for the disease. They contrast with the rare high penetrance cancer susceptibility genes, like *TP53, BRAC1, BRAC2* etc, seen in other cancers. No such high penetrance susceptibility genes have been found in CLL[12].

But, because the effect of individual SNP variants is weak, they cannot be used individually for predicting the risk of CLL. However an ingenious Polygenic Risk Score (PRS) was recently proposed integrating the information of 41 representative SNP variants[16], as has been done in other cancers. In this way the authors found a strong correlation of a higher PRS with risk of CLL and of monoclonal B-cell lymphocytosis (MBL), which was independent of the effect of age and family history (FH). The PRS therefore does not account for all the FH risk seen in familial CLL but may be additive[16]. In other respects this observation may not have clinical application but it confirms the model of the common risk alleles as involved in the pathogenesis of the disease.

Monoclonal B-cell lymphoytosis (MBL)

It has now been clearly established that MBL represents a precursor condition for CLL. Some patients with MBL have normal lymphocyte counts ($<4 \times 10^9$/L) and others have moderate lymphocytosis (total lymphocyte count $\geq 4 \times 10^9$/L). By definition MBL requires $<5 \times 10^9$/L monoclonal B lymphocytes in order to distinguish it from CLL. Patients with MBL share cytogenetic abnormalities with CLL patients, chiefly 13q14 del, trisomy 12 and skewed *IGHV* usage. 15% will go on to develop CLL. Monoclonal B-lymphocytes with the immunophenotype of CLL can be found in the peripheral blood of c.3.5% of adult individuals[17,18].

An important observation in relation to familial CLL is that MBL was detected in 13.5% of "healthy" first-degree relatives from 21 CLL families, with a significantly higher incidence than in the general population. As in the latter, the proportion increases with age[19]. The detection of MBL could therefore become a surrogate marker for the genetic predisposition to CLL. Another study[20] confirmed that the incidence of MBL was increased among relatives of CLL families with two or more affected members.

MBL was a feature of the individual (seen in **Figure 1**) who sparked my interest in familial CLL in the 1980s.

In order to explore this further a GWAS study was carried out to ascertain whether the 10 variant SNPs, known at the time to confer a risk of CLL could also predispose to MBL[21]. An association between genotype and MBL was found for 9 out of the 10 SNPs. Six of them were statistically significant[21], especially the SNPs nearest to the genes *IRF4* and *GRAMD1B*[21].

The relevance of the GWAS findings in MBL was recently confirmed and extended by means of the PRS combining 41 representative SNPs[16]. This PRS may help identify those individuals at greater risk of developing MBL and CLL.

In addition to the similar risk alleles, identified in GWAS studies, MBL is indistinguishable from CLL at the somatic level (tested in the monoclonal B cells), although the burden of driver mutations is lower than in fully blown CLL. This is consistent with the concept that the evolution from MBL to CLL is accomplished by the progressive accumulation of driver mutations[22].

Germline mutations in shelterin complex genes

Despite the advances in the identification of risk alleles in CLL, so far no specific gene mutations have been described in familial CLL. Only rare germline alterations (in the normal cells of patients with CLL) were reported in the *ATM* (ataxia telangiectasia mutated) gene[23-25], associated with CLL risk. These rare *ATM* germline mutations were not found in familial cases and are associated with somatic mutations/deletions (in the leukemia cells) in the other allele, therefore suggesting that *ATM* may act as a tumor suppressor gene[25].

The availability of next generation sequencing technology allowed our group to undertake WES in 66 CLL families with 141 affected individuals and to report for the first time germline mutations in *POT1* (protection of telomeres 1) and two other genes of the shelterin complex: *TERF21P* and *ACD*. The initial search was undertaken in order to identify loss of function mutations, prioritising those genes with a known role in cancer and excluding the more common variants. Sanger sequencing was subsequently performed for confirmation[11].

A total of seven families were found with the same germline mutation in each affected individual (four with *POT1* and three with the other genes), suggesting an incidence of 6% in familial CLL. In a complementary analysis of 1083 unselected CLL cases and 5854 controls, one of the variant *POT1* mutations, which has a global allele frequency of 0.0005, was found in 6 cases and seems to confer a 3.81-fold increased risk of CLL[11].

POT1 encodes a component of the shelterin complex genes, part of which binds directly to DNA, and it is the first member of this telomere structure to be found mutated in a human cancer[26].

The relevance of *POT1* in CLL is threefold: i) it is mutated in 3.5% of CLL cases (by somatic mutation) and causes telomere dysfunction[26]; ii) one of the 42 SNPs associated with CLL risk is near *POT1*[27]; and now iii) germline mutations in *POT1* have been found to be involved in families co-segregating CLL[11,28]. The latter findings are also consistent with earlier evidence of linkage of familial CLL to chromosomes 7q31.32-q33 and 16q12.2-q23[10].

Shelterin is a telomere specific protein complex composed of six family members (*POT1*, etc) which are essential for all telomere functions, including protection from degradation. Of note, germline mutations in *POT1* have also been found in familial cutaneous melanoma and glioma but none of the familial CLL cases with shelterin mutations had either of those conditions[11].

What does it all mean?

There have been major advances in our understanding of the genetic architecture of CLL in the last decade and this has been facilitated by the high familial component, the highest familial risk among all the lymphoid malignancies. Rare high penetrance cancer susceptibility genes have not been found in CLL like in some cancers, e.g. *BRCA1* and *BRAC2* in breast cancer. The major factor of risk is the existence of common low-penetrance risk alleles whose function is currently being elucidated and highlighted by the GWAS studies[13,16]. The demonstration of germline mutations in the shelterin complex genes in seven families suggests that their effect may be of moderate penetrance[11,12]. In addition *POT1* and the other shelterin genes mutated in the germline of CLL families are sitting at the unique interface between the somatic driver mutations in sporadic CLL, of which 80 so far have been identified[22,28,29],

and the 42 susceptibility loci in sporadic and familial CLL[13,28].

As illustrated in a masterly editorial in Blood[28], commenting on the shelterin paper[11], the only other gene mutated at the germline in rare cases is *ATM,* but not in familial CLL. *BCL2* and *IRF4* are also at the interface, but either as susceptibility SNP loci or as somatic mutations[28]. It is not clear how the variant SNPs, most of them in non-coding regions of the genome, influence CLL. However, many tend to map to areas of active chromatin and seem to be implicated in dysregulation of immunity genes and B-cell development[13].

Declaración de conflictos de interés:
El autor declara que no posee conflictos de interés:

References

1. Yuille MR, Matutes E, Marossy A, Hilditch B, Catovsky D, Houlston RS. Familial chronic lymphocytic leukaemia: a survey and review of published studies. Br J Haematol. 2000;109:794-799.

2. Houlston RS, Catovsky D, Yuille MR. Genetic susceptibility to chronic lymphocytic leukemia. Leukemia. 2002;16:1008-1014.

3. Videbaek A. Heredity in Human Leukemia and Its Relation to Cancer: a Genetic and Clinical Study of 209 Probands. 1947. HK Lewis and Co. Ltd., London.

4. Videbaek A. Familial leukemia. Acta Path Microbiol Scand 1958;44:372.

5. Gunz FW, Veale AM. Leukemia in close relatives--accident or predisposition? J Natl Cancer Inst. 1969;42:517-524.

6. Gunz FW, Gunz JP, Veale AM, Chapman CJ, Houston IB. Familial leukaemia: a study of 909 families. Scand J Haematol. 1975;15:117-131.

7. Jønsson V, Houlston RS, Catovsky D y col. CLL family 'Pedigree 14' revisited: 1947-2004. Leukemia. 2005;19:1025-1028.

8. 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;104:1850-1854.

9. Crowther-Swanepoel D, Wild R, Sellick G y col. Insight into the pathogenesis of chronic lymphocytic leukemia (CLL) through analysis of IgVH gene usage and mutation status in familial CLL. Blood. 2008;111:5691-5693.

10. Sellick GS, Goldin LR, Wild RW y col. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. Blood. 2007;110:3326-3333.

11. Speedy HE, Kinnersley B, Chubb D y col. Germ line mutations in shelterin complex genes are associated with familial chronic lymphocytic leukemia. Blood. 2016;128:2319-2326.

12. Sud A, Kinnersley B, Houlston RS.Genome-wide association studies of cancer: current insights and future perspectives. Nat Rev Cancer. 2017;17:692-704.

13. Law PJ, Berndt SI, Speedy HE y col. Genome-wide association analysis implicates dysregulation of immunity genes in chronic lymphocytic leukaemia. Nat Commun. 2017;8:14175.

14. Di Bernardo MC, Crowther-Swanepoel D, Broderick P y col. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2008;40:1204-1210.

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

16. Kleinstern G, Camp NJ, Goldin LR y col. Association of polygenic risk score with the risk of chronic lymphocytic leukemia and monoclonal B-cell lymphocytosis. Blood. 2018;131:2541-2551.

17. Rawstron AC, Green MJ, Kuzmicki A y col. Monoclonal B lymphocytes with the characteristics of "indolent" chronic lymphocytic leukemia are present in 3.5% of adults with normal blood counts. Blood. 2002;100:635-639.

18. Rawstron AC, Bennett FL, O'Connor SJ y col. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. N Engl J Med. 2008;359:575-583.

19. Rawstron AC, Yuille MR, Fuller J y col. Inherited predisposition to CLL is detectable as subclinical monoclonal B-lymphocyte expansion. Blood. 2002;100:2289-2290.

20. Goldin LR, Lanasa MC, Slager SL y col. Common occurrence of monoclonal B-cell lymphocytosis among members of high-risk CLL families. Br J Haematol. 2010;151:152-158.

21. Crowther-Swanepoel D, Corre T, Lloyd A y col. Inherited genetic susceptibility to monoclonal B-cell lymphocytosis. Blood. 2010;116:5957-5960.

22. Puente XS, Beà S, Valdés-Mas R y col. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015;526:519-524.

23. Stankovic T, Weber P, Stewart G y col. Inactivation of ataxia telangiectasia mutated gene in B-cell chronic lymphocytic leukaemia. Lancet. 1999;353:26-29.

24. Yuille MR, Condie A, Hudson CD y col. ATM mutations are rare in familial chronic lymphocytic leukaemia. Blood. 2002;100:603-609.

25. Tiao G, Improgo MR, Kasar S y col. Rare germline variants in ATM are associated with chronic lymphocytic leukemia. Leukemia. 2017;31:2244-2247.

26. Ramsay AJ, Quesada V, Foronda M y col. POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. Nat Genet. 2013;45:526-530.

27. Speedy HE, Di Bernardo MC, Sava GP y col. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2014;46:56-60.

28. Beekman R, Campo E. Shelterins, a genetic crossroad in CLL. Blood. 2016;128:2279-2280.

29. Landau DA, Carter SL, Stojanov P y col. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell. 2013;152:714-726.