

Brief report

The IGHV1-69/IGHJ3 recombinations of unmutated CLL are distinct from those of normal B cells

Francesco Forconi,^{1,2} Kathleen N. Potter,³ Elisa Sozzi,¹ Isla Henderson,³ Emanuele Cencini,¹ Davide Rossi,⁴ Riccardo Bomben,⁵ Valter Gattei,⁵ Gianluca Gaidano,⁴ Graham Packham,² and Freda K. Stevenson³

¹Hematology, Department of Clinical Medicine and Immunological Sciences, University of Siena and Department of Oncology, Azienda Ospedaliera Universitaria Senese, Siena, Italy; ²Cancer Sciences Unit, Cancer Research UK Centre, University of Southampton Faculty of Medicine, Southampton, United Kingdom; ³Genetic Vaccine Group, Cancer Sciences Unit, University of Southampton Faculty of Medicine, Southampton, United Kingdom; ⁴Division of Hematology, Department of Clinical and Experimental Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; and ⁵Clinical and Experimental Onco-Hematology Unit, Centro di Riferimento Oncologico, Istituto di Ricovero e Cura a Carattere Scientifico, Aviano, Italy

IGHV1-69/51p1 is expressed by ~ 30% of unmutated chronic lymphocytic leukemia (U-CLL) and combines with selected IGHD and IGHJ genes generating stereotypes if HCDR3 amino acid homology is > 60%. We had previously revealed stereotypic IGHV1-69/IGHJ6 rearrangements in normal naive B cells, thereby identifying potential counterparts of U-CLL. A different stereotypic IGHV1-69/IGHD3-16(RF2)/IGHJ3 rearrangement carrying the CAR(GGx)YD

motif in the N1-region, recurrent in 6% IGHV1-69+ve CLL, is exceptionally sequence restricted, strongly suggestive of shared antigen recognition. We have now analyzed IGHV1-69/IGHJ3 rearrangements in circulating B cells of healthy individuals using several PCR-based approaches with IGHV1-69/IGHJ3 CLL sequences for reference. Stereotypes were found, but all were distinct from CLL. Remarkably, even a highly sensitive

semi-nested PCR, specific for the CLL-expressed IGHV1-69/IGHD3-16(RF2)/IGHJ3 stereotype, failed to identify the CAR(GGx)YD sequence, although similar motifs were found. These highly specific B cells are not apparent in the accessible normal repertoire and may expand in response to rarely expressed antigens important in the pathogenesis of CLL. (Blood. 2012;119(9):2106-2109)

Introduction

The IGHV1-69/51p1 allele is expressed by ~ 30% of unmutated chronic lymphocytic leukemia (U-CLL).¹⁻³ In these cases, IGHV1-69 rearranges to a restricted number of IGHD and IGHJ, generating conserved sequences in the third complementarity-determining regions (HCDR3). Using a 60% amino-acid (AA) homology cut-off, ~ 50% of cases can be assigned to stereotypic subsets suggestive of shared antigen recognition.²

The B cell of origin of CLL has long been sought and these conserved sequences provide tools for probing the normal adult expressed repertoire. We previously revealed a significant proportion (~ 33%) of stereotypes involving IGHV1-69/IGHJ6 rearrangements among normal circulating naive B cells, several (~ 20%) being similar to those of U-CLL.⁴ However, these incorporate the long IGHJ6, which accounts for a large proportion of the similarity.^{2,5,6} A more restricted HCDR3 is found in the CLL subset 6,² where IGHV1-69, combined with IGHD3-16 in reading-frame 2 (RF2) with the shorter IGHJ3, generates almost identical sequences between cases. This is highly suggestive of shared antigen recognition,^{2,7} especially as there is unusual conservation of the nontemplated N-regions, generating the characteristic CAR(GGx)YD N1 and SYR(xND)AFDI N2 junctional motifs.⁷⁻⁹ Frequent combination with the same IGKV3-20/A27 adds strength to this conclusion.⁷ Here we have probed IGHV1-69/IGHJ3 rearrangements in the circulating B cells of healthy donors, specifically seeking counterparts of subset 6.

Methods

Amplification of IGHV1-69/51p1-related rearrangements from healthy donors

IGHV1-69/51p1-related rearrangements were amplified, cloned and sequenced from peripheral blood lymphocyte cDNA of 3 healthy individuals,^{4,10} in 2 centers. First, the forward IGHV1-69/HCDR2 primer⁴ was used with C μ 100 or IGHJ3 primers (prJ3, prDJ3, prNDJ3) with different levels of extension into the CLL-specific N2-region (supplemental Table 1, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article). The IGHJ3 primers were designed to provide a hierarchy from no (prJ3) to high specificity (prNDJ3) for subset 6-related N2-region containing the asparagine-aspartic acid (ND) motif.^{7,8} Second, a semi-nested PCR/cloning approach was used to detect subset 6 sequences in donors. This included a first PCR with forward IGHV1-69/HCDR2 primer and IGHJ3 reverse primers prDJ3 or prNDJ3 followed by a second round with the reverse prs6D3-16 primer specific for subset 6 (supplemental Figure 1). This approach increased the sensitivity of detection of subset 6-related rearrangements to < 1 in 10⁷ B cells (supplemental Figure 2).

Analysis of IGHV1-69/IGHJ3 combinations, HCDR3 clustering and search for subset 6 IGHV1-69/IGHD3-16(RF2)/IGHJ3 rearrangements in healthy donors

A CLL database (n = 653) of 256 previously published,⁴ and 397 new IGHV1-69-D-J rearrangements centralized at 4 different institutions was

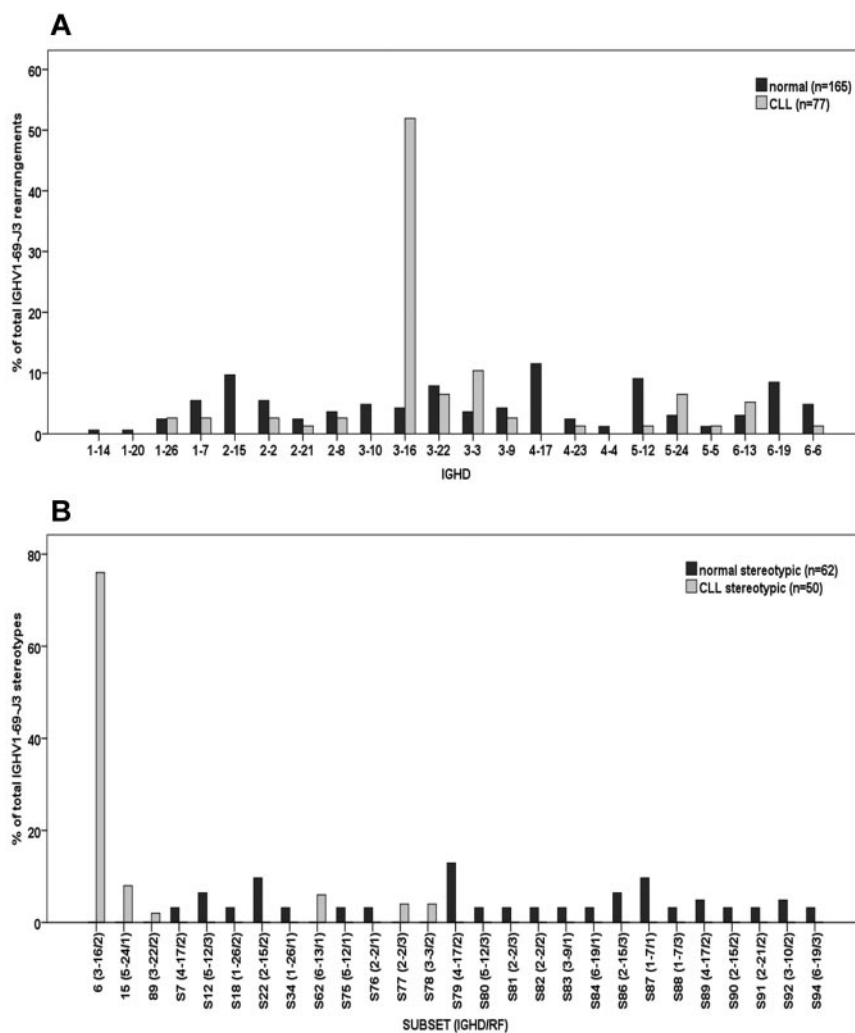
Submitted August 23, 2011; accepted December 27, 2011. Prepublished online as *Blood* First Edition paper, January 10, 2012; DOI 10.1182/blood-2011-08-375501.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2012 by The American Society of Hematology

Figure 1. IGHD and Stereotype distribution in normal and CLL B cells using IGHV1-69/IGHJ3 rearrangements. Bars identify percent sequences (y-axis) using IGHD (A) or assigned to each subset (B) of all normal (dark gray) or CLL (light gray) IGHV1-69/IGHJ3 rearrangements. (A) Percent IGHD distribution among normal and CLL B cells refers to all IGHV1-69/IGHJ3 rearrangements analyzed (stereotypic and nonstereotypic). (B) Percent stereotypic pattern distribution refers to those normal and CLL B cells using IGHV1-69/IGHJ3 rearrangements assigned to subsets only. For each subset, IGHD gene and reading frame are also indicated (in parentheses separated by a slash) in the x-axis.



used for comparative sequence analysis. First, we compared the 77 CLL derived rearrangements using IGHJ3 (supplemental Table 2), with those from the healthy donors. All sequences were aligned to the ImMunoGenetics (IMGT) directory and analyzed for IGHV1-69 mutational status, IGHD, IGHJ, HCDR3 junctional characteristics, and stereotypes.^{2,3,6,11} Identical IGHD-J usage and IGHD reading-frame (RF), ClustalW2 score > 60, AA identity > 60% and < 3 AA difference in length were required, exactly as published.^{2-4,12} Numbering of subsets was updated according to previous criteria.²⁻⁴ Second, 65 CLL subset 6 sequences (see supplemental Table 2 and publications)^{2,3,5,7,8,12-20} were compared with those obtained from the sensitive semi-nested PCR/cloning procedures used to extract subset 6 sequences from healthy individuals.

The study was approved by IRBs at the Universities of Southampton (228/02/t) and Siena (AIRC IG-5298, ITT-2008). Patients and donors provided informed consent before inclusion in the study in accordance with the Declaration of Helsinki.

Results and discussion

IGHV1-69 recombinations in immunoglobulin M (IgM)-positive normal blood B cells involve a range of IGHJ genes with IGHJ3 accounting for ~ 20%,¹⁰ and this was confirmed as 16.8% (23/137) in the present study (supplemental Table 3). The frequency in the current CLL database was 11.8% (77/653; supplemental Table 4), as found previously.^{2,4,10} From normal B cells, we obtained 23 IGHV1-69/IGHJ3 sequences using the IGHV1-69 and C μ 100

primer pair. To increase the number of sequences expressing IGHJ3, we took advantage of our previous observation that B cells expressing IGHV1-69, detectable with the G6 monoclonal antibody, were within the naive IgM-positive circulating population.⁴ We therefore moved to a IGHJ3-specific primer, confident that amplicons were derived from these B cells. This allowed collection of 165 IGHV1-69/IGHJ3 rearrangements (supplemental Table 5) for comparison with the 77 from CLL.

Almost all normal B cell sequences were unmutated (163/165, 98.8%) as were most from CLL (65/77, 84.4%). Cases assigned to subset 15 (CLL only) or to no subset (either normal or CLL) were the exceptions.² However a striking difference in IGHD usage was evident, with normal B cells expressing a wide range, but CLL cells predominantly (40/77, 51.95%) expressing IGHD3-16 (Figure 1A). Stereotypic sequences defined by pre-established criteria, including 60% homology,^{2-4,12} were evident in normal B cells (62/165, 37.6%) and, as expected from IGHD-gene usage, were wide-ranging (Figure 1B). In contrast, CLL stereotypes reflected the preferential usage of IGHD3-16(RF2), with 38/77, 49.4% (76% of total stereotypes; Figure 1B) in subset 6, and 12/77 assigned to other subsets. Strikingly, none of the stereotypic sequences from normal B cells clustered with any CLL subsets (Figure 1B). Among the few normal cells using IGHD3-16 (7/165, 4.24%), RF2 was detected in 5 (71.4%), but none of the sequence patterns fulfilled the strict pre-established criteria for assignment to subset 6.^{2-4,12}

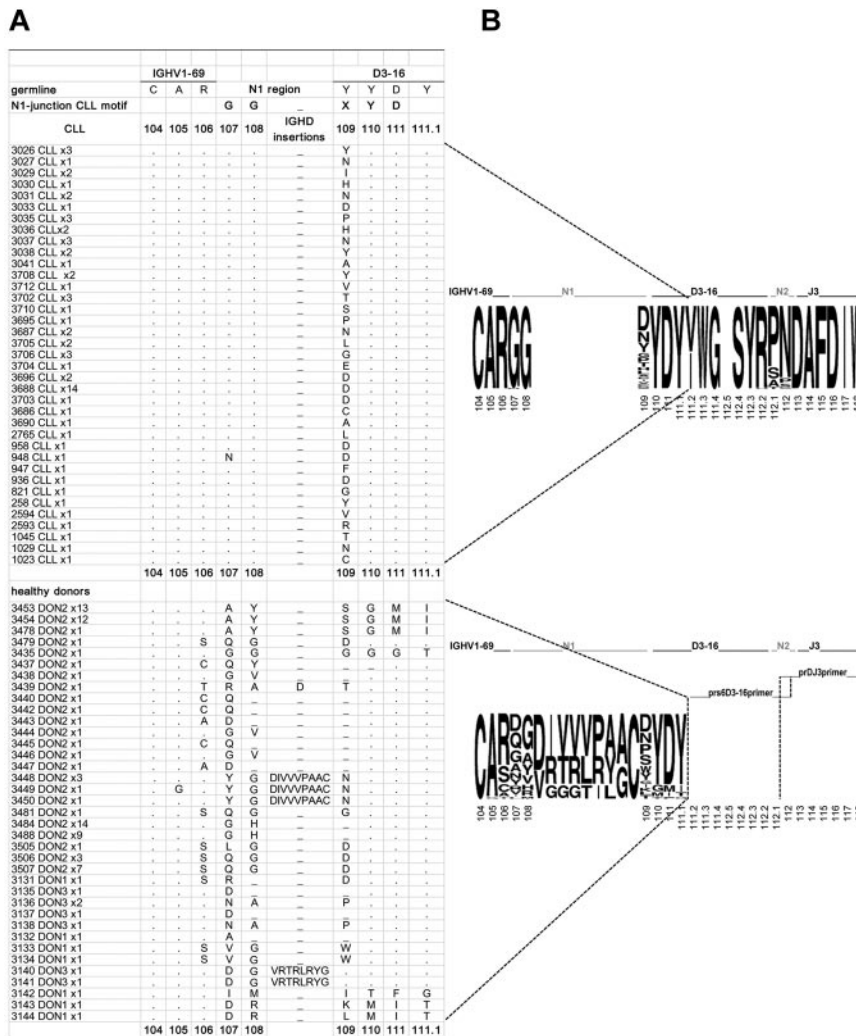


Figure 2. Amino acid sequence alignments of HCDR3 region from CLL-derived IGH rearrangements assigned to subset 6 and from healthy donor-derived IGH rearrangements identified with a semi-nested subset 6 specific primer approach. (A) Left column: the first number indicates sequence identification number (SIN.); DON1/2/3 indicate donor's source; x followed by a number (eg, x13) indicates the number of identical sequences (eg, 13 identical sequences) identified among different tumor rearrangements from CLL individuals or by cloning of PCR products from normal individuals. (.) indicates homology; and (-), lack of amino acid. Among CLL sequences represented with identical (GGx)YD motif in separate lines, different amino acids were observed at codon 111.2 and/or N2 junctional region (codons 112.1-112). Among normal sequences, a different nucleotide sequence at codon 111.2, leading different amino acid translation outside primer sequence was observed between clones 3448 and 3450. Among the healthy donors, those amino acid sequences, identical from position 107 to 111.1 but different outside the represented amino acid sequence or in the degenerate prs6D3-16 primer sequence at position 111.2 in IGHJ3-16 and/or 112-113 in the N2 junctional region, are indicated with different SIN. (B) Crude percent frequency by logo analysis representation. Top row: alignment for CLL subset 6 rearrangements. Bottom row: normal sequences identified by cloning after PCR amplification with HCDR2-J3 and subsequent semi-nested HCDR2-prs6D3-16 primers approach. Amino acid sequences are represented as sequence logos as described at <http://weblogo.berkeley.edu/>. The letters represent the amino acids used at each particular codon numbered according to IMGT criteria. When more than one amino acid was observed in a codon, the letters representing each change are displayed as a stack. The size of the amino acid symbol represents the crude percent frequency of any of the observed amino acid in that position. Sequences with no amino acid at a given position were excluded from the calculation of the frequency at that codon. Blank spaces represent absence of any amino acid in that particular position as a comparison between CLL and normal sequences. Codon numbering between IMGT positions 108 and 109 is omitted.

To probe more specifically among normal B cells for the characteristic conserved HCDR3 sequence of subset 6, a primer was designed to move further into the IGHJ3-16/N2/IGHJ3 junction (supplemental Table 1 and supplemental Figure 1). A single-step PCR ($> 10^{-5}$ sensitivity) yielded no products, but semi-nested PCR with pr(N)DJ3 and prs6D3-16 primers ($> 10^{-7}$ sensitivity) generated amplicons. We identified 37 sequences with unique HCDR3s together with multiple repetitive sequences, possibly reflecting limited available template (Figure 2 and supplemental Table 6). Among the wide range of HCDR3 sequences in normal B cells (Figure 2), the subset 6 defining motif CAR(GGx)YD (codons 107-111 counting from first G),^{7,8} characteristic of CLL sequences, where X varies (Figure 2), was not found. In contrast to the 3 AA (GGx) of the N1 junctional region, normal B cells displayed variable lengths ranging from 3-12 AA (Figure 2). Variability was because of single AA deletions/insertions and/or insertion of IGHJ3-2 (sequences 3448-50) or another unknown segment (sequences 3140-1) upstream of IGHJ3-16.

Further probing revealed several sequences involving IGHJ3-16/IGHJ3 which, however, did not meet the strict subset 6-specific criteria of identity to the CAR(GGx)YD motif. The closest sequences were CAS(QG_)YD (Clone 3506), CAS(QGx)YD (3481) CAS(QGx)YD (3479) and CAS(VGx)YD (3431-34) that all have

2-3 AA in N1, including at least 1 G residue and variable X, as for CLL subset 6 (Figure 2).

This analysis of IGHV1-69/IGHJ3 recombinants in the normal naive B-cell repertoire reveals heterogeneity of HCDR3 sequences, as expected in innate B cells programmed to recognize a wide range of common pathogens. Similar heterogeneity was observed in the IGHV1-69/IGHJ6 recombinants.⁴ However, the range of these unmutated sequences appears relatively limited, possibly reflecting the low affinity and consequent cross-reactivity of antigen binding.²¹ This limitation leads to stereotypes, subsequently reflected in transformed B cells of CLL. The unusually restricted HCDR3 of IGHV1-69/IGHJ3-16(RF2)/IGHJ3 (subset 6) represents one end of the spectrum. The fact that we found similar, albeit not identical, sequences again indicates the limited range of the repertoire. The reason for some expansion of this subset in CLL is likely to involve antigen drive, but one possibility is that the sequences have to interact both with the initiating pathogen and with an autoantigen. The nonmuscle myosin heavy chain IIA antigen exposed on apoptotic bodies is recognized by CLL subset 6 immunoglobulins,⁸ and may promote expansion and survival of CLL cells. Similar cross-reactivity has been proposed for other CLL derived immunoglobulins,²² and has potential implications for driving proliferation in vivo and therefore clinical outcome.

Acknowledgments

The authors thank Victoria Bryant for input to sequence analysis. They especially thank Dr Kostas Stamatopoulos (Democritus University of Thrace, Alexandroupolis, Greece) and Dr Nikos Darzentas (Institute of Agrobiotechnology, Thessaloniki, Greece) for the help with stereotypic analysis.

This study was supported by Cancer Research UK (CR-UK) and the Southampton CR-UK Center; Experimental Cancer Medicine Center, Tenovus Solentside, Associazione Italiana per la Ricerca sul Cancro (AIRC IG-5298 and Special Program Molecular Clinical Oncology, 5 × 1000, No. 10007), Milan, Italy; Istituto Toscano Tumori, Italy; Siena-AIL (Associazione Italiana contro le Leucemie, Linfomi e Mielomi, Siena Section); Hairy Cell Leukemia Research Foundation (Illinois); Fondazione MPS 2009 (Siena, Italy); Ministry of University and Research-Project of national interest 2009 (PRIN-MUR 2009).

Authorship

Contribution: F.F. designed the study, interpreted data, performed analysis, and wrote the manuscript; K.N.P. designed the study,

interpreted data, performed analysis, and contributed to writing the manuscript; E.S., E.C., and I.H. performed research and IGHV analysis and interpreted data; R.B. and D.R. provided well-characterized biologic samples and molecular and clinical data of the reference cohort; V.G., G.G., and G.P. contributed to interpretation of the data; and F.K.S. designed the study, interpreted data, and wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Dr Francesco Forconi, Cancer Research UK Centre, Cancer Sciences Unit, University of Southampton Faculty of Medicine, Southampton University Hospital Trust, Tremona Road, SO16 6YD, Southampton, United Kingdom; e-mail: f.forconi@soton.ac.uk; or Hematology, Department of Clinical Medicine and Immunological Sciences, University of Siena & Department of Oncology, AOUS, Viale Bracci, 16, 53100, Siena, Italy; e-mail: forconif@unisi.it; or Prof Freda Stevenson, Genetic Vaccine Group, Cancer Sciences Division, Southampton University Hospitals Trust, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, United Kingdom; e-mail: fs@soton.ac.uk.

References

- Kipps TJ, Tomhave E, Pratt LF, Duffy S, Chen PP, Carson DA. Developmentally restricted immunoglobulin heavy chain variable region gene expressed at high frequency in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A*. 1989;86(15):5913-5917.
- Murray F, Darzentas N, Hadzidimitriou A, et al. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood*. 2008;111(3):1524-1533.
- Stamatopoulos K, Belessi C, Moreno C, et al. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood*. 2007;109(1):259-270.
- Forconi F, Potter KN, Wheatley I, et al. The normal IGHV1-69-derived B-cell repertoire contains stereotypic patterns characteristic of unmutated CLL. *Blood*. 2010;115(1):71-77.
- Johnson TA, Rassenti LZ, Kipps TJ. Ig VH1 genes expressed in B cell chronic lymphocytic leukemia exhibit distinctive molecular features. *J Immunol*. 1997;158(1):235-246.
- Lefranc MP, Giudicelli V, Kaas Q, et al. IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res*. 2005;33(database issue):D593-D597.
- Widhopf GF, 2nd Rassenti LZ, Toy TL, Gribben JG, Wierda WG, Kipps TJ. Chronic lymphocytic leukemia B cells of more than 1% of patients express virtually identical immunoglobulins. *Blood*. 2004;104(8):2499-2504.
- Chu CC, Catera R, Hatzl K, et al. Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood*. 2008;112(13):5122-5129.
- Chu CC, Catera R, Zhang L, et al. Many chronic lymphocytic leukemia antibodies recognize apoptotic cells with exposed nonmuscle myosin heavy chain IIA: implications for patient outcome and cell of origin. *Blood*. 2010;115(19):3907-3915.
- Potter KN, Orchard J, Critchley E, Mockridge CI, Jose A, Stevenson FK. Features of the overexpressed V1-69 genes in the unmutated subset of chronic lymphocytic leukemia are distinct from those in the healthy elderly repertoire. *Blood*. 2003;101(8):3082-3084.
- 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.
- Messmer BT, Albesiano E, Efremov DG, et al. Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J Exp Med*. 2004;200(4):519-525.
- Gurrieri C, McGuire P, Zan H, et al. Chronic lymphocytic leukemia B cells can undergo somatic hypermutation and intraclonal immunoglobulin V(H)DJ(H) gene diversification. *J Exp Med*. 2002;196(5):629-639.
- Nikitin EA, Pivnik AV, Sudarikov AB, et al. [A comparison of the forms of chronic lymphocytic leukemia in relation to the mutational status of the genes of the immunoglobulin variable region]. *Ter Arkh*. 2000;72(7):52-56.
- Nardini E, Neri F, Vicenzi E, et al. Thymic function and immunoglobulin mutation genotype in B-cell chronic lymphocytic leukemia patients. *Int J Cancer*. 2003;107(6):958-961.
- Tschumper RC, Geyer SM, Campbell ME, et al. Immunoglobulin diversity gene usage predicts unfavorable outcome in a subset of chronic lymphocytic leukemia patients. *J Clin Invest*. 2008;118(1):306-315.
- Lanemo Myhrinder A, Hellqvist E, Sidorova E, et al. A new perspective: molecular motifs on oxidized LDL, apoptotic cells, and bacteria are targets for chronic lymphocytic leukemia antibodies. *Blood*. 2008;111(7):3838-3848.
- Stamatopoulos K, Belessi C, Hadzidimitriou A, et al. Immunoglobulin light chain repertoire in chronic lymphocytic leukemia. *Blood*. 2005;106(10):3575-3583.
- Tobin G, Thunberg U, Karlsson K, et al. Subsets with restricted immunoglobulin gene rearrangement features indicate a role for antigen selection in the development of chronic lymphocytic leukemia. *Blood*. 2004;104(9):2879-2885.
- Giudicelli V, Chaume D, Lefranc MP. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res*. 2005;33(database issue):D256-D261.
- Herve M, Xu K, Ng YS, et al. Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. *J Clin Invest*. 2005;115(6):1636-1643.
- Rosen A, Murray F, Evaldsson C, Rosenquist R. Antigens in chronic lymphocytic leukemia—implications for cell origin and leukemogenesis. *Semin Cancer Biol*. 2010;20(6):400-409.