Comprehensive characterization of IGHV3-21–expressing B-cell chronic lymphocytic leukemia: an Italian multicenter study

Riccardo Bomben,1 Michele Dal Bo,1 Daniela Capello,2 Dania Benedetti,1 Daniela Marconi,3 Antonella Zucchetto,1 Francesco Forconi,4 Rossana Maffei,5 Emanuela M. Ghia,6 Luca Laurenti,7 Pietro Bulian,1 Maria Ilaria Del Principe,8 Giuseppe Palermo,3 Mia Thorsellius,9 Massimo Degan,1 Renato Campanini,3 Anna Guarini,5 Giovanni Del Poeta,8 Richard Rosenquist,9 Dimitar G. Efremov,10 Roberto Marasca,9 Robin Foà,6 Gianluca Gaidano,2 and Valter Gattei1

1Clinical and Experimental Hematology Research Unit, Centro di Riferimento Oncologico, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Aviano (PN, Italy); 2Division of Hematology, Department of Clinical and Experimental Medicine and Interdisciplinary Research Center on Autoimmune Diseases (IRCAD), Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; 3Department of Physics, University of Bologna, Italy; 4Division of Hematology and Transplant, Department of Clinical Medicine and Immunological Sciences, University of Siena, Italy; 5Division of Hematology, Department of Oncology and Hematology, University of Modena and Reggio Emilia, Modena, Italy; 6Division of Hematology, Department of Cellular Biotechnologies and Hematology, University La Sapienza, Rome, Italy; 7Hematology Institute, Catholic University Sacro Cuore, Rome, Italy; 8Chair of Hematology, S Eugenio Hospital and University of Tor Vergata, Rome, Italy; 9Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Sweden; 10International Centre for Genetic Engineering and Biotechnology (ICGEB) Outstation-Monterotondo, Consiglio Nazionale delle Ricerche (CNR) Campus A. Buzzati-Traverso, Rome, Italy.

IGHV3-21–using chronic lymphocytic leukemia (CLL) is a distinct entity with restricted immunoglobulin gene features and poor prognosis and is more frequently encountered in Northern than Southern Europe. To further investigate this subset and its geographic distribution in the context of a country (Italy) with both continental and Mediterranean areas, 37 IGHV3-21 CLLs were collected out of 1076 cases enrolled by different institutions from Northern or Central Southern Italy. Of the 37 cases, 18 were identified as homologous (hom)HCDR3–IGHV3-21 CLLs and were found almost exclusively (16 of 18) in Northern Italy; in contrast, 19 nonhomHCDR3–IGHV3-21 cases were evenly distributed throughout Italy. Clinically, poor survivals were documented for IGHV3-21 CLLs as well as for subgroups of mutated and homHCDR3–IGHV3-21 CLLs. Negative prognosticators CD38, ZAP-70, CD49d, and CD79b were expressed at higher levels in homHCDR3 than nonhomHCDR3–IGHV3-21 cases. Differential gene expression profiling (GEP) of 13 IGHV3-21 versus 52 non–IGHV3-21 CLLs identified, among 122 best-correlated genes, TGFβ2 and VPR1 as down- and up-regulated in IGHV3-21 CLL cases, respectively. Moreover, GEP of 7 homHCDR3 versus 6 nonhomHCDR3–IGHV3-21 CLLs yielded 20 differentially expressed genes, with WNT-16 being that expressed at the highest levels in homHCDR3–IGHV3-21 CLLs. Altogether, IGHV3-21 CLLs, including those with homHCDR3, had a peculiar global phenotype in part explaining their worse clinical outcome. (Blood. 2007; 109:2989-2998)

© 2007 by The American Society of Hematology

Introduction

The clinical course of B-cell chronic lymphocytic leukemia (CLL) can be in part foreseen by the presence of mutated (M) or unmutated (UM) immunoglobulin variable (IGV) genes.1 Analysis of IGV heavy chain (IGHV) gene usage has revealed expression of a biased IGV repertoire in CLL compared with normal B cells.3,4 Moreover, a significant fraction of CLL expresses B-cell receptors (BCRs) characterized by nonrandom stereotyped combinations of IGV-D-J rearrangements often associated with a restricted selection of IGV light chains.3-6 Such molecular peculiarities occur both as homologous (hom)HCDR3–IGHV3-21 CLLs and were found almost exclusively (16 of 18) in Northern Italy; in contrast, 19 nonhomHCDR3–IGHV3-21 cases were evenly distributed throughout Italy. Clinically, poor survivals were documented for IGHV3-21 CLLs as well as for subgroups of mutated and homHCDR3–IGHV3-21 CLLs. Negative prognosticators CD38, ZAP-70, CD49d, and CD79b were expressed at higher levels in homHCDR3 than nonhomHCDR3–IGHV3-21 cases. Differential gene expression profiling (GEP) of 13 IGHV3-21 versus 52 non–IGHV3-21 CLLs identified, among 122 best-correlated genes, TGFβ2 and VPR1 as down- and up-regulated in IGHV3-21 CLL cases, respectively. Moreover, GEP of 7 homHCDR3 versus 6 nonhomHCDR3–IGHV3-21 CLLs yielded 20 differentially expressed genes, with WNT-16 being that expressed at the highest levels in homHCDR3–IGHV3-21 CLLs. Altogether, IGHV3-21 CLLs, including those with homHCDR3, had a peculiar global phenotype in part explaining their worse clinical outcome. (Blood. 2007; 109:2989-2998)

© 2007 by The American Society of Hematology

The online version of this article contains a data supplement.
biased geographic distribution and to provide a comprehensive biologic and clinical characterization of IGHV3-21 CLL, we investigated cases derived from 1076 unselected CLL patients enrolled by different reference centers located in Northern (N) or Central Southern (CS) Italy. Because IGHV3-21 CLL patients have a clinical behavior independent of the IGHV gene mutation status, we investigated the differential gene expression profile characterizing IGHV3-21 CLL versus non–IGHV3-21 cases. The resulting profile and the function of some differentially regulated genes integrated the information provided by the unique study published so far. Finally, we were able to identify a number of genes characterizing the differential expression signature of IGHV3-21 CLL with homologous versus nonhomologous HCDR3.

**Patients, materials, and methods**

**Patients, healthy donors, and human cell lines**

The study included peripheral blood (PB) samples from 1076 unselected CLL patients collected from the archives of the Clinical and Experimental Hematology Research Unit of the Aviano National Cancer Institute (n = 136); Division of Hematology of the Amedeo Avogadro University of Eastern Piedmont, Novara (n = 110); Department of Oncology and Hema-

tology of the University of Modena-Reggio Emilia (n = 284); Hematology Unit of La Sapienza University, Rome (n = 228); S Eugenio Hospital, University of Tor Vergata, Rome (n = 127); Hematology Institute of the Catholic University, Rome (n = 95); and the Hematology Unit of the University of Siena (n = 96). According to the location of each center, 2 main geographic subgroups were identified: the N Italy subgroup, which included 530 patients recruited by the Aviano, Novara, and Modena-Reggio Emilia centers (N patients), and the CS Italy subgroup, with 546 patients recruited by the Siena center and the institutions located in Rome (CS patients) (Table 1). Median ages were 64.0 and 60.4 years, while male-female ratios were 1.6 and 1.4 for N and CS patients, respectively.

Normal B cells used in gene expression profiling (GEP) experiments were purified from PB samples from 13 healthy donors kindly provided by the Blood Bank of the Aviano National Cancer Institute.

For survival studies, we combined 3 single-institution series (Aviano, 69 patients; Modena-Reggio Emilia, 152; Roma Tor Vergata, 80; total, 301) in which the standard cutoff of 2% IGHV mutations proved to have prognostic value in the context of each single series (not shown); the overall mean follow-up of the 301 patients was 66 months (range, 3 to 289). Patient survival was analyzed using Kaplan-Meier survival curves and log-rank test. Comparisons of frequencies between specific CLL subsets were made by applying the Pearson χ² test with Yates continuity correction.

Sources, phenotypic characterization, and culture conditions of human cell lines (THP1, KG1, HEL, HUT102, CEM, Granita-519, LP1, L1236) have been reported in detail previously;2,16,20; the Hodgkin disease–derived cell line L1236 was a gift from V. Diehl (University of Cologne, Germany).

**Analysis of IGHV, IGKV, and IGLV rearrangements and HCDR3 clustering**

IGHV, IGKV kappa (IGKV), and IGLV rearrangements were amplified from either reverse-transcribed total RNA or genomic DNA, as previously described.2,16 Purified amplicons were sequenced either directly or upon insertion into plasmid vectors and cloning.2,7,8 Sequences were aligned to the IgBLAST, VBASE, or ImMunoGeneTics directories and considered mutated if deviation from the corresponding germ line gene was at least 2%. For IGHD gene determination, a requirement of at least 7 matching nucleotides was used. HCDR3-driven clustering was performed by converting all in-frame IGHV rearrangements into amino acid (aa) sequences and by aligning HCDR3 sequences by means of the multiple sequence alignment software ClustalW.21 The length of HCDR3 was computed between codons 95 and 102, as described by Kabat.22 For consistency with previous studies,12,15,16,17 IGVL chains were also indicated according to the Kawasaki nomenclature.

**Flow cytometry**

All immunophenotypical determinations were performed at the Aviano National Cancer Institute. Sources, specificities, and fluorochrome combinations of monoclonal antibodies along with details on the working protocols were reported elsewhere.23 Detection of ZAP-70 was performed as reported using a cutoff point of 30% to define positivity.24,25

**B-cell purification, RNA extraction, and gene expression profiling (GEP) analysis**

Neoplastic and normal B cells (more than 95% pure), purified by Ficoll-Hypaque gradient (PharmaCia, Upsala, Sweden) and positive immunoselection with anti-CD19–conjugated immunomagnetic beads (Miltenyi Biotec, Bologna, Italy), were employed as total RNA source. cDNA synthesis, cDNA purification, in vitro transcription of amino-allyl RNA (aRNA), aRNA dye coupling, and purification of dye-labeled aRNA were performed using the Amino Allyl MessageAmp II aRNA Amplification Kit (Ambion, Austin, TX) following manufacturer’s guidelines. Cy3/Cy5 was from Amersham Biosciences (Amersham, United Kingdom). Cy3/Cy5 dye incorporation into aRNA yielded incorporation rates of 30 to 60 dye molecules per 1000 nucleotides by spectrophotometric analysis, as requested by the manufacturer.

GEP was performed by a dual labeling strategy using Cy3-labeled aRNA from pooled normal PB cells of healthy donors as common reference and Cy5-labeled aRNA from purified CLL cells as tester. A mixture of 4 μg Cy3-labeled reference aRNA and 4 μg Cy5-labeled aRNA from each CLL was hybridized to Operon Human Genome Oligo Set Version 2.0 platforms (21,329 oligomers of 70 nucleotides; Operon Biotechnologies, Huntsville, AL) and spotted in duplicates on MicroMax SuperChip I glass slides (Perkin Elmer Life Sciences, Boston, MA). Hybridization was performed for 18 hours at 48°C in a buffer containing 5× standard sodium citrate (SSC), 0.1% sodium dodecyl sulfate (SDS), 25% formamide, and 100 μg/mL salmon sperm DNA using an automated HybArray 12 hybridization system (Perkin Elmer Life Sciences). Following washings, glass slides were analyzed by a Scan Array Lite scanner (Packard BioChips Technologies, Billerica, MA); images and data were analyzed using the GenePix Pro software (Axon Instruments, Foster City, CA). The final data set was analyzed by means of unsupervised and supervised clustering as detailed in Supplemental Materials and Methods, “Bioinformatics approach for analysis of gene expression profiling (GEP) of IGHV3-21 CLLS” (available on the Blood website; see the Supplemental Materials link at the top of the online article).

**Reverse transcriptase–polymerase chain reaction (RT-PCR) and quantitative-real-time RT-PCR (QRT-PCR)**

Expression of TGFβ2, VIPR1, and WNT-16 was evaluated by both conventional and quantitative (Q) RT-PCR using, in the latter case, SYBR Green dye–containing reaction buffer (Applied Biosystems, Foster City, CA). Reactions were performed in the presence of 0.5 pmol/μL of primers specific for TGFβ2 (sense: 5′-AGAGAGCGCAGCAAGGTTCA-3′; anti
sense: 5′-AATTCTGAGTGGATGCAATAA-3′), VIPR1 (sense: 5′-GGCTGGTGCGTGGTGCTGG-3′; antisense: 5′-GACCAGGAGACCTTCGCTTG-3′), and WNT-16 (sense: 5′-CGGGGACCAATGAGACCAACA-3′; antisense: 5′-GAGGAGCAGTTCCAGAACCA-3′). Amplification protocol was 4 minutes (10 minutes for QRT-PCR) at 95°C for initial denaturation, 35 cycles (40 for QRT-PCR) of 15 seconds at 95°C, 30 seconds at 62°C (57°C for TGFβ2), and 90 seconds at 72°C, followed by final extension at 72°C for 5 minutes. Incorporation of SYBR Green dye into the PCR products was monitored in real time with an ABI PRISM 7700 sequence detection system (Applied Biosystems). Standard curves (10-fold dilution from 10⁰ to 10⁻⁴ attomoles) were prepared for both target genes and the housekeeping gene β2-microglobulin (βM). For each experimental sample, the amount of target and endogenous reference was determined from the appropriate standard curve, and the target amount was divided by the endogenous reference amount to
Results

Distribution of IGHV families in CLL from N and CS Italy

Of 1076 patients with CLL entering this study, 530 and 546 patients were enrolled by institutions located in Northern (N) and Central Southern (CS) Italy, respectively. As summarized in Figure 1, IGHV gene families with overall similar frequencies in N and CS patients. In agreement with other smaller series,10,27,28 the IGHV3 family was most frequently used (N, 49%; CS, 52%), followed by IGHV1 (N, 21%; CS, 21%) and IGHV4 (N, 24%; CS, 20%) (Figure 1). Considering the 2% cutoff to discriminate between M and UM status, 2% of CLLs (39.2%) were classified as UM, equally distributed among N (313 cases, 59.1%) and CS (341 cases, 62.5%). Similarly, 422 CLLs (39.2%) were classified as UM, equally distributed among N (313 cases, 59.1%) and CS (341 cases, 65.4). Of 1076 patients with CLL entering this study, 530 and 546 patients were enrolled by institutions located in Northern (N) and Central Southern (CS) Italy, respectively. As summarized in Figure 1, IGHV gene family usage was analyzed in CLL from Northern (530 cases) (A) and Central Southern Italy (546 cases) (B). Frequencies were calculated for the whole series (●) and according to the IGHV mutational status (UM, □; M, □)

A given IGHV segment was recognized in 28 of 37 IGHV3-21 cases (7.4%), all displaying an in-frame and potentially functional rearrangement. Table 2 summarizes the characteristics of IGHV3-21 rearrangements in our series. Twenty-five of 37 cases (67.6%) expressed an IGHV gene with more than 2% mutations, with a median mutation rate of 4.0% (range, 2.1% to 7.1%); 20 of 25 M IGHV3-21 CLLs had a "low" mutation burden of 2% to 5%.

A given IGHV segment was recognized in 28 of 37 IGHV3-21 rearrangements. IGHV5-24 and IGHV3-22 were the most common segments, detected in 4 rearrangements each. A specific IGHV segment was not assigned in 9 cases, 8 of them expressing the highly homologous short HCDR3 sequences described in detail below. Finally, the IGH6 gene was the most frequent IGHJ segment (20 of 37 cases), followed by IGHJ4 (9 of 37 cases).

Cluster analysis of putative translated HCDR3 protein sequences allowed us to split IGHV3-21 CLL into 2 groups (Table 2; Figure 2). The first group (18 of 37 cases) displayed a highly homologous and short HCDR3 sequence (homHCDR3–IGHV3-21 CLL) characterized by the aa sequence DANGMDV in 6 of 18 cases (Figure 2). The remaining cases expressed a similar 7-aa stretch modified by 1 (6 cases) or 2 (5 cases) aa's; a single case (Mo107) had 3 aa changes in the conserved 7-aa sequence (Figure 2). Analyses of nucleotide sequences of homHCDR3–IGHV3-21 CLL often (8 of 18 cases) yielded an unidentified IGHD segment always associated with IGH6 gene usage (Table 2). The second subset of IGHV3-21 CLLs (19 of 37 cases) expressed HCDR3 sequences usually longer (median length, 14 nucleotides; range, 11 to 23 nucleotides) and totally unrelated to each other (nonhom-HCDR3–IGHV3-21 CLLs; Figure 2). Expression of longer HCDR3 sequences from this group may explain why identification of IGHV3-21 segments was possible in all but 1 case (Table 2).

We then investigated whether the 7-aa stretch expressed by homHCDR3–IGHV3-21 CLL was also employed by IGHV rearrangements other than IGHV3-21. HCDR3 sequences from IGHV3-21 CLLs were clustered along with HCDR3 sequences from 349 CLL cases representative of the whole series (Figure S1). Results of cluster analysis clearly identified a cluster that included all homHCDR3–IGHV3-21 CLLs along with 1 single non–IGHV3-21 case (case B383 rearranging IGHV3-66; Figure S1), suggesting that use of homologous HCDR3 characterized by the 7-aa stretch is very rare in non–IGHV3-21 CLL.

Table 1. Characteristics of CLL patients

<table>
<thead>
<tr>
<th></th>
<th>Av</th>
<th>No</th>
<th>Mo-RE</th>
<th>Total</th>
<th>Si</th>
<th>RoLS</th>
<th>RoTV</th>
<th>RoCU</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>M; UM</td>
<td>76</td>
<td>60</td>
<td>63</td>
<td>47</td>
<td>313</td>
<td>237</td>
<td>110</td>
<td>91</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>136</td>
<td>284</td>
<td>530</td>
<td></td>
<td>25</td>
<td>141</td>
<td>87</td>
<td>127</td>
<td>95</td>
</tr>
</tbody>
</table>

A total of 1076 unselected patients were enrolled by institutions located either in Northern (N patients) or Central Southern (CS patients) Italy. The institutions located in Northern Italy were the Aviano National Cancer Institute (Av), the Division of Hematology of the Novara University of Eastern Piedmont (No), and the Department of Oncology and Hematology of the University of Modena-Reggio Emilia (Mo-RE); the institutions located in Central Southern Italy were the Hematology Unit of the University of Siena (Si), the Hematology Unit of the La Sapienza University of Rome (RoLS), the S Eugenio Hospital of the University of Tor Vergata in Rome (RoTV), and the Hematology Institute of the Catholic University in Rome (RoCU).

M indicates mutated IGHV gene status; UM, unmutated IGHV gene status.
Table 2. General characteristics of heavy chain and light chain rearrangements in IGHV3-21 CLLs

<table>
<thead>
<tr>
<th>CLL case</th>
<th>IGHV</th>
<th>IGHD</th>
<th>IGHJ</th>
<th>% mutation*</th>
<th>Similar HCDR3, y/n†</th>
<th>IGKV(IGL) and IGKV-J genes</th>
<th>% mutation‡</th>
<th>Similar L(K)CDR3, y/n†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mo66_N</td>
<td>3-21</td>
<td>1-1</td>
<td>6</td>
<td>0.6</td>
<td>y</td>
<td>IGLV3-21</td>
<td>2.0</td>
<td>y</td>
</tr>
<tr>
<td>Mo107_N</td>
<td>3-21</td>
<td>1-22</td>
<td>6</td>
<td>1.2</td>
<td>y</td>
<td>IGLV3-21 &amp; IGLV1-51</td>
<td>4.6 &amp; 3.4</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>B15_N</td>
<td>3-21</td>
<td>5-18</td>
<td>6</td>
<td>1.4</td>
<td>y</td>
<td>IGLV3-21 &amp; IGLV1-40</td>
<td>1.6 &amp; 2.9</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>No4937_N</td>
<td>3-21</td>
<td>na</td>
<td>6</td>
<td>1.5</td>
<td>y</td>
<td>IGLV3-21</td>
<td>0.0</td>
<td>y</td>
</tr>
<tr>
<td>B281_N</td>
<td>3-21</td>
<td>5-18</td>
<td>6</td>
<td>2.1</td>
<td>y</td>
<td>IGLV3-21 &amp; IGLV1-40</td>
<td>1.9 &amp; 4.1</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>Mo91_N</td>
<td>3-21</td>
<td>1-26</td>
<td>6</td>
<td>2.3</td>
<td>y</td>
<td>IGLV3-21 &amp; IGLV4-3</td>
<td>1.2 &amp; 2.9</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>No480_N</td>
<td>3-21</td>
<td>na</td>
<td>6</td>
<td>2.4</td>
<td>y</td>
<td>IGLV3-21</td>
<td>1.7</td>
<td>y</td>
</tr>
<tr>
<td>Mo124_N</td>
<td>3-21</td>
<td>na</td>
<td>6</td>
<td>2.9</td>
<td>y</td>
<td>IGLV3-21</td>
<td>1.2 &amp; 1.2</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>Mo145_N</td>
<td>3-21</td>
<td>na</td>
<td>6</td>
<td>3.0</td>
<td>y</td>
<td>IGLV3-21 &amp; IGLV1-51</td>
<td>1.2 &amp; 5.3</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>Mo105_N</td>
<td>3-21</td>
<td>na</td>
<td>6</td>
<td>3.4</td>
<td>y</td>
<td>IGLV3-21</td>
<td>0.4</td>
<td>y</td>
</tr>
<tr>
<td>B304_N</td>
<td>3-21</td>
<td>1-14</td>
<td>6</td>
<td>3.5</td>
<td>y</td>
<td>IGLV3-21 &amp; IGLV1-51</td>
<td>2.4 &amp; 4.8</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>B189_N</td>
<td>3-21</td>
<td>na</td>
<td>6</td>
<td>3.8</td>
<td>y</td>
<td>IGLV3-21</td>
<td>2.6</td>
<td>y</td>
</tr>
<tr>
<td>No4844_N</td>
<td>3-21</td>
<td>5-24</td>
<td>6</td>
<td>3.9</td>
<td>y</td>
<td>IGLV3-21</td>
<td>2.1</td>
<td>y</td>
</tr>
<tr>
<td>Mo2_N</td>
<td>3-21</td>
<td>na</td>
<td>6</td>
<td>4.0</td>
<td>y</td>
<td>IGLV3-21 &amp; IGLV2-8</td>
<td>1.7 &amp; 2.6</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>Mo50_N</td>
<td>3-21</td>
<td>na</td>
<td>6</td>
<td>4.2</td>
<td>y</td>
<td>IGLV3-21 &amp; IGLV6-57</td>
<td>1.6 &amp; 4.0</td>
<td>y &amp; n</td>
</tr>
<tr>
<td>B164_N</td>
<td>3-21</td>
<td>5-24</td>
<td>6</td>
<td>4.4</td>
<td>y</td>
<td>IGLV3-21</td>
<td>1.5</td>
<td>y</td>
</tr>
<tr>
<td>B205_CS</td>
<td>3-21</td>
<td>5-24</td>
<td>6</td>
<td>0.4</td>
<td>y</td>
<td>IGLV3-21</td>
<td>1.9</td>
<td>y</td>
</tr>
<tr>
<td>B292_CS</td>
<td>3-21</td>
<td>5-24</td>
<td>6</td>
<td>1.4</td>
<td>y</td>
<td>IGLV3-21</td>
<td>2.0</td>
<td>y</td>
</tr>
<tr>
<td>No1139_N</td>
<td>3-21</td>
<td>3-10</td>
<td>6</td>
<td>0.0</td>
<td>n</td>
<td>IGKV2-30</td>
<td>0.0</td>
<td>n</td>
</tr>
<tr>
<td>Mo1_N</td>
<td>3-21</td>
<td>3-22</td>
<td>4</td>
<td>0.0</td>
<td>n</td>
<td>IGKV2-28 &amp; IGKV3-20</td>
<td>0.0 &amp; 2.9</td>
<td>n</td>
</tr>
<tr>
<td>Mo28_N</td>
<td>3-21</td>
<td>3-10</td>
<td>4</td>
<td>0.0</td>
<td>n</td>
<td>IGLV3-12</td>
<td>0.8</td>
<td>n</td>
</tr>
<tr>
<td>Mo37_N</td>
<td>3-21</td>
<td>6-13</td>
<td>4</td>
<td>1.2</td>
<td>n</td>
<td>IGLV9-49</td>
<td>0.0</td>
<td>n</td>
</tr>
<tr>
<td>B110_N</td>
<td>3-21</td>
<td>3-22</td>
<td>4</td>
<td>3.0</td>
<td>n</td>
<td>IGKV1-39</td>
<td>0.6</td>
<td>n</td>
</tr>
<tr>
<td>Mo158_N</td>
<td>3-21</td>
<td>6-13</td>
<td>6</td>
<td>4.8</td>
<td>n</td>
<td>IGKV4-1</td>
<td>1.3</td>
<td>n</td>
</tr>
<tr>
<td>B193_N</td>
<td>3-21</td>
<td>4-17</td>
<td>4</td>
<td>7.0</td>
<td>n</td>
<td>IGLV3-21</td>
<td>2.3</td>
<td>y</td>
</tr>
<tr>
<td>S3827_CS</td>
<td>3-21</td>
<td>3-22</td>
<td>3</td>
<td>0.0</td>
<td>n</td>
<td>IGLV3-21</td>
<td>0.0</td>
<td>y</td>
</tr>
<tr>
<td>S3569_CS</td>
<td>3-21</td>
<td>3-10</td>
<td>4</td>
<td>0.7</td>
<td>n</td>
<td>IGKV4-1</td>
<td>0.0</td>
<td>n</td>
</tr>
<tr>
<td>B338_CS</td>
<td>3-21</td>
<td>3-10</td>
<td>5</td>
<td>2.5</td>
<td>n</td>
<td>IGKV1D-8</td>
<td>2.9</td>
<td>n</td>
</tr>
<tr>
<td>G124_CS</td>
<td>3-21</td>
<td>7-27</td>
<td>4</td>
<td>3.8</td>
<td>n</td>
<td>IGKV1D-8</td>
<td>2.9</td>
<td>n</td>
</tr>
<tr>
<td>B321_CS</td>
<td>3-21</td>
<td>6-13</td>
<td>4</td>
<td>4.4</td>
<td>n</td>
<td>IGKV1-8</td>
<td>1.2</td>
<td>n</td>
</tr>
<tr>
<td>S3667_CS</td>
<td>3-21</td>
<td>6-19</td>
<td>3</td>
<td>4.4</td>
<td>n</td>
<td>IGKV1D-8</td>
<td>2.1</td>
<td>n</td>
</tr>
<tr>
<td>S5153_CS</td>
<td>3-21</td>
<td>3-22</td>
<td>6</td>
<td>4.4</td>
<td>n</td>
<td>IGKV1-17</td>
<td>1.5</td>
<td>n</td>
</tr>
<tr>
<td>S5427_CS</td>
<td>3-21</td>
<td>3-22</td>
<td>4</td>
<td>5.1</td>
<td>n</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>S45_CS</td>
<td>3-21</td>
<td>6-13</td>
<td>2</td>
<td>6.5</td>
<td>n</td>
<td>IGKV2-14</td>
<td>3.3</td>
<td>n</td>
</tr>
<tr>
<td>G125_CS</td>
<td>3-21</td>
<td>6-6</td>
<td>4</td>
<td>6.8</td>
<td>n</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>B305_CS</td>
<td>3-21</td>
<td>na</td>
<td>5</td>
<td>7.1</td>
<td>n</td>
<td>IGKV1-12</td>
<td>6.0</td>
<td>n</td>
</tr>
</tbody>
</table>

The heavy and light chain rearrangements were performed by using either reverse-transcribed total RNA or genomic DNA. Purified amplicons were either directly sequenced or sequenced upon insertion into plasmid vector and cloning. In the latter case, percent mutation is the mean value of the percent mutations found in each clone. Each case is identified by a unique patient number; letters reported after the underscore symbol identify the Northern (N) or Central Southern (CS) geographic areas.

*Refers to IGHV mutational status.
†Evidence (y) or not (n) of similar aa HCDR3 or L(K)CDR3 composition.
‡Refers to IGL/JV mutational status.
§With the only exclusion of light chain data of case Mo2, data reported for these cases are exactly as reported by Thorselius et al.17
†Evidence (y) or not (n) of similar aa HCDR3 or L(K)CDR3 composition.
*Refers to IGHV mutational status.

Molecular features of IGHV3-21 CLL: IGKV and IGLV rearrangements

Among 34 IGHV3-21 cases with available IGV light chain rearrangement data, a total of 44 rearrangements, including 32 IGL and 12 IGK, were found (Table 2). HomHCDR3–IGHV3-21 CLL showed the IGLV3-21(V,2-14)/IGLJ3 rearrangement either alone (9 cases) or in association with another potentially functional IGLV rearrangement (9 cases). These rearrangements included another IGLV3-21/V,2-14 (1 case), IGLV1-51/V,1-19 (3 cases), IGLV1-40/V,1-13 (2 cases) or, in the remaining cases, IGLV5-45/V,5-1, IGLV2-8/V,1-2, or IGLV6-57/V,1-22 (Table 2). Conversely, nonhomHCDR3–IGHV3-21 cases rearranged either IGLV (5 cases) or IGKV (11 cases) without any bias for specific genes. Two potentially functional IGV light chain rearrangements were amplified in a single case (Mo1).

Distribution of IGHV3-21 CLL in N and CS Italy

As summarized in Table 3, the percentage of IGHV3-21 CLls was slightly higher in N (23 of 530, 4.3%) versus CS patients (14 of 546, 2.6%), although this difference did not reach statistical significance (P = .15). However, most homHCDR3–IGHV3-21 cases (16 of 18) were recruited by institutions located in N Italy. In particular, homHCDR3–IGHV3-21 CLls represented 3.0% (16 of 530) of N cases and only 0.4% (2 of 546) of CS cases (P = .002; Table 3). A statistical significance (P = .003) was also reached when comparing the sole distribution of homHCDR3 and nonhomHCDR3 IGHV3-21 cases in the 2 main geographic areas (data not shown). Conversely, distribution of nonhomHCDR3–IGHV3-21 CLls was not dissimilar in N and CS patient series (P = .39; Table 3).
Survival data were analyzed in 301 patients, 279 with a non–IGHV3-21 CLL (176 with M and 103 with UM IGHV genes) and 22 with IGHV3-21 CLL (14 M and 8 UM). This latter group included 15 homHCDR3 and 7 nonhomHCDR3–IGHV3-21 CLLs. In the context of non–IGHV3-21 CLL (Figure 3A), a significantly longer survival was documented for M cases (median survival, not reached; 95% confidence interval [CI], 173 to not reached) compared with UM CLLs (median survival, 104 months; 95% CI, 90 to not reached). The survival curve of IGHV3-21 CLL (median survival, 154 months; 95% CI, 93 to not reached) was similar to that of UM non–IGHV3-21 cases ($P = 0.312$) and significantly different from M CLLs ($P = 0.030$; Figure 3A). An analogous trend was observed when only homHCDR3–IGHV3-21 CLLs were considered (Figure 3B). Moreover, both M IGHV3-21 CLL and M homHCDR3–IGHV3-21 CLLs displayed significantly shorter survivals when compared with M non–IGHV3-21 cases (Figure 3C-D). Finally, no differences were documented by comparing survival of UM (8 cases) versus M (14 cases) IGHV3-21 CLL as well as nonhomHCDR3– (7 cases) versus homHCDR3–IGHV3-21 CLL (15 cases) (data not shown).

Expression of immunophenotypic prognosticators by IGHV3-21 CLL

Expression data for the negative immunophenotypic prognosticators CD38, ZAP-70,25,29-32 CD49d, and CD79b 33,34 were available for 14 IGHV3-21 CLLs (8 homHCDR3 and 6 nonhomHCDR3–IGHV3-21 cases; 10 cases with M and 4 cases with UM IGHV genes). HomHCDR3–IGHV3-21 cases more frequently expressed these negative prognosticators above the cutoff value of 30% of positive cells23,24,34 as compared with nonhomHCDR3–IGHV3-21 CLLs ($P = 0.03$) (Figure 4A). Moreover, while 3 of 6 nonhomHCDR3–IGHV3-21 CLLs expressed none of the negative prognosticators, all the homHCDR3–IGHV3-21 cases expressed at least 2 (7 of 8 cases) or 1 (all cases) negative prognosticators (Figure 4B). Of note, all 4 UM CLLs (cases B15, B205, B292 and No4937 in Figure 4B) had a homHCDR3–IGHV3-21 sequence and, with the only exclusion of CD79b for B292 and CD38/CD49d for No4937, expressed all the negative prognosticators above the established threshold (Figure 4B).

### Table 3. Distribution of IGHV3-21 in the main Italian geographic areas

<table>
<thead>
<tr>
<th>Geographic Area</th>
<th>$\text{IGHV3-21 B-CLLs with homHCDR3}^*$</th>
<th>$\text{IGHV3-21 B-CLLs with nonhomHCDR3}^*$</th>
<th>$\text{IGHV3-21 B-CLLs with homHCDR3}^+$</th>
<th>$\text{IGHV3-21 B-CLLs with nonhomHCDR3}^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Italy</td>
<td>16/530 (.002†)</td>
<td>7/530 (.39†)</td>
<td>23/530 (.15†)</td>
<td>14/530 (—)</td>
</tr>
<tr>
<td>Central Southern Italy</td>
<td>2/246 (—)</td>
<td>12/246 (—)</td>
<td>14/246 (—)</td>
<td>246 (—)</td>
</tr>
</tbody>
</table>

$*$Number of cases in a given CLL subgroup over the total number of CLLs in the specific geographic area.

†$P$ values obtained by the Pearson $\chi^2$ test with Yates continuity correction, comparing Northern versus Central Southern Italy CLL subgroups.
Differential gene expression profiling (GEP) of IGHV3-21 CLL

Thirteen IGHV3-21 (11 M and 2 UM) and 52 non–IGHV3-21 CLLs (35 M and 17 UM) were analyzed for GEP. An unsupervised approach was used in exploratory analyses to evaluate whether IGHV3-21 and non–IGHV3-21 CLLs could be separated by a hierarchic clustering (ie, had a distinguishable global GEP). This unsupervised analysis was performed on a total of 28 513 features, which passed the prefiltering steps aimed at eliminating all features that were either not expressed or overall not differentially modulated, as detailed in Supplemental Materials and Methods, “Bioinformatics approach for analysis of gene expression profiling (GEP) of IGHV3-21 CLLs.” The generated hierarchic clustering (Figure S2) clearly demonstrated that all 13 IGHV3-21 CLLs were intermingled among non–IGHV3-21 cases, thus indicating that the 2 subgroups shared the large part of the features investigated.

Therefore, to discover genes differentially expressed in IGHV3-21 versus non–IGHV3-21 CLL, we applied a supervised approach capable of detecting minimal differences in gene expression of otherwise related cell populations. In particular, to address this issue in the context of an unbalanced data set (13 IGHV3-21 versus 52 non–IGHV3-21 cases), we used an original modified version of significance analysis of microarrays (SAM) statistical analysis with an adjusted $P \leq 10^{-2}$, yielded a list of 20 differentially expressed features (Table S3). Among these features, 12 were down-regulated and 8 up-regulated in homHCDR3–IGHV3-21 CLLs; in this latter group, the gene of known function with the highest score was $WNT-16$ (Table S3; Figure 5B). Real-time PCR experiments confirmed high $WNT-16$ expression in representative homHCDR3–IGHV3-21 CLL as well as in the Granta-519 B-cell line (Figure 5C).

Discussion

In the present report, we provide a comprehensive characterization of 37 IGHV3-21 CLLs collected out of a series of 1076 unsellected cases enrolled by different reference centers located in N or CS Italy. The study addressed the geographic distribution of IGHV3-21 CLL in the context of different areas of the same country (ie, Italy) as well as the molecular, immunophenotypical, and clinical features of this CLL subset. The peculiar clinical behavior of IGHV3-21 CLL patients prompted us to investigate whether either IGHV3-21 or homHCDR3–IGHV3-21 CLLs were characterized by specific GEP explaining, at least in part, the disease’s aggressiveness.
Figure 5. IGHV3-21–associated gene expression profiles in CLL. (A) Heat map generated with 122 probes found to be differentially expressed between IGHV3-21 patients (red bar under the tree) and non–IGHV3-21 patients (blue bar) by supervised analyses (Supplemental Materials). The color scale identifies relative gene expression changes normalized by the SD of 1 with 0 representing the mean expression level of a given gene. Expression for the 2 probes for the IGLV3-21(V/H9261 2-14) gene is reported below the heat map (location of these probes in the heat map is indicated by dashes); results of IGV light chain rearrangements, when available (ie, for all IGHV3-21 cases and 1 non–IGHV3-21 CLL), are reported below the corresponding hybridization spot. Location in the heat map of the VIPR1 and TGFβ2 genes are indicated by 1 and 2 asterisks, respectively. (B) Heat map generated with 20 genes found to be differentially expressed between homHCDR3–IGHV3-21 cases (yellow bar under the tree) and nonhomHCDR3–IGHV3-21 cases (orange bar) by supervised analyses (see “Materials and methods”). Location in the heat map of the WNT-16 gene is indicated by 3 asterisks. (C) Validation of microarray results by QRT-PCR. Three representative genes were selected to validate microarray results by QRT-PCR: VIPR1 and TGFβ2 were chosen to characterize non–IGHV3-21 versus IGHV3-21 CLLs (13 total cases), while WNT-16 was chosen among the genes differentially expressed by homHCDR3–IGHV3-21 and nonhomHCDR3–IGHV3-21 cases (6 cases). The incorporation of the SYBR Green dye into the PCR products was monitored in real time, and the resulting threshold cycles (Ct) (ie, the PCR cycle number corresponding to the beginning of the exponential growth of PCR products) were computed. Ct values were converted into attomoles by means of QRT-PCR experiments carried out with serial dilution of known concentrations of VIPR1-, TGFβ2-, WNT-16-, and βM-specific amplicons. Results for each CLL case represent the values of the relative expression levels of VIPR1(βM) and TGFβ2(β2M) in IGHV3-21 cases (red bars) and non–IGHV3-21 cases (blue bars) and WNT-16(β2M) in homHCDR3–IGHV3-21 cases (yellow bars) and nonhomHCDR3–IGHV3-21 VH3-21 cases (orange bars). The panel on the bottom right represents ethidium bromide–stained agarose gels showing the expression of VIPR1, TGFβ2, and WNT-16 by conventional RT-PCR in 8 cell lines and in 3 representative CLL cases (B338, B80, and B419). Asterisks refer to molecular weight markers.
Biased geographic distribution of IGHV3-21 CLL in Italy

An overrepresentation (about 10%) of IGHV3-21 CLL has been documented in several Northern European countries. Conversely, a multicenter study involving countries from Southern Europe, as well as studies on separate cohorts of patients, collectively indicated a significantly lower percentage of IGHV3-21 cases, ranging from 0% to about 3%. Notably, our results point to heterogeneous distribution of homHCDR3–IGHV3-21 CLL between continental and Mediterranean areas of the same country (ie, Italy). Indeed, 16 of 18 (89%) homHCDR3–IGHV3-21 CLls were enrolled by centers located in N Italy. Therefore, homHCDR3–IGHV3-21 CLL accounted for 3.0% (16 of 530) in N Italy but was virtually absent in CS Italy (2 of 546, 0.4%). A referral bias among N versus CS centers seems unlikely because (1) the frequency of IGHV3-21 usage by CLL cells overall was similar in N and CS patients and (2) the M/U M ratio (1.5 in the total series) was similar in N and CS cohorts of patients (1.44 and 1.66, respectively).

The incidence of homHCDR3–IGHV3-21 CLL in a smaller Italian series was 3 of 97 (3.1%) in this regard, it seems likely that the Italian series reported by the multicentric study of Southern European countries contains an unwanted bias, because such series was allegedly composed of patients recruited in N Italy. Curiously, the percentage of homHCDR3–IGHV3-21 CLls detected by our study in CS Italy (0.4%) overlaps with that found in other Mediterranean areas, namely Greece and Spain. The reasons for the different incidence of homHCDR3–IGHV3-21 CLL in N versus CS Italy or, more in general, in typically Mediterranean areas versus continental regions are currently unclear.

Molecular characterization of IGHV3-21 CLL

As expected, almost 50% of IGHV3-21 cases displayed an IGHV gene rearrangement with a very short and homologous HCDR3 segment, often expressing the DANGMDV aa motif. HCDR3 cluster analysis documented that this 7-aa motif is very conserved in IGHV3-21 CLls and expressed by homHCDR3–IGHV3-21 CLls detected by our study in CS Italy (0.4%) but was not found in other Mediterranean areas, namely Greece and Spain. The reasons for the different incidence of homHCDR3–IGHV3-21 CLL in N versus CS Italy or, more in general, in typically Mediterranean areas versus continental regions are currently unclear.

Clinical features and expression of immunophenotypic prognosticators in IGHV3-21 CLL

In agreement with previous studies evaluating the clinical impact of IGHV3-21 CLL versus non–IGHV3-21 cases, the survival of IGHV3-21 CLL in our series was comparable with that of UM patients. Differences in overall survival also held true when only the 15 homHCDR3–IGHV3-21 CLL cases were considered or when the 14 M IGHV3-21 cases and the 10 M homHCDR3–IGHV3-21 CLls were compared with their M non–IGHV3-21 counterparts. Collectively, these findings suggest that we should withdraw IGHV3-21 cases when evaluating the prognostic impact of IGHV gene mutations in CLls.

According to our data on overall survival, we failed to confirm the worse prognosis of homHCDR3–IGHV3-21 cases compared with nonhomHCDR3–IGHV3-21 CLls suggested by one group but not by another. However, our results indicate that a set of well-established negative immunophenotypic prognosticators are more frequently expressed by homHCDR3–IGHV3-21 CLL than by nonhomHCDR3–IGHV3-21 cases. The low number of total cases, along with the presence of 4 of 8 UM CLls in the homHCDR3–IGHV3-21 subset, prevents definitive conclusions on this matter. Notably, the finding of a differential gene expression profile distinguishing homHCDR3–IGHV3-21 from nonhom–IGHV3-21 CLls, characterized by the overexpression of genes allegedly enhancing the aggressiveness of CLL phenotype, could be in keeping with the association of homHCDR3–IGHV3-21 CLL with a poor-prognosis phenotype (see “GEP of IGHV3-21 CLL”).

GEP of IGHV3-21 CLL

Using an unsupervised approach, IGHV3-21 and non–IGHV3-21 CLL shared the large part of 28513 features investigated. This finding is not surprising, because a similar approach was unable to split CLL into discrete categories. Nevertheless, subtle GEP differences of phenotypically related cell populations can be detected using supervised analyses. The multi-SAM approach used in this study had the additional advantage of overcoming the problem of an unbalanced representation of IGHV3-21 CLL (n = 13) versus non–IGHV3-21 CLL (n = 52) in our data set. The probes eventually selected for correlation with IGHV3-21 usage were 122, corresponding to 115 genes. As an internal control, IGLV3-21(V,2-14) fell among the best-ranked genes up-regulated in IGHV3-21 CLL, while probes related to IGHV3-21 were not available in the employed platform.

Genes down-regulated in IGHV3-21 CLL included 4 negative regulators of cell proliferation (RERG, ABI1, PMP22, TGFβ2). In particular, TGFβ2, whose production by CLL has been previously demonstrated, had expression levels significantly lower in IGHV3-21 CLL cells by both GEP and RTQ-PCR validation experiments. Noteworthily, the down-regulated expression in IGHV3-21 CLL cells of 2 downstream molecules of the TGFβ signaling cascade (eg, MADH2/SMAD2 and AML1/RUNX1) has been demonstrated by Falt et al and in this study, respectively. In addition, the gene coding for the TGFβ receptor I was selected as down-regulated in IGHV3-21 CLL by multi-SAM, although not at a ranking as high as to be included in the 122 best-correlated probes. Furthermore, the regulatory subunit 15B (PP1R15B) of protein phosphatase 1 (PP1), a gene inhibiting the TGFβ signaling pathway, was significantly up-regulated in IGHV3-21 CLL and included in the 122 best-correlated probes. Taken together, present data and data reported by Falt et al collectively demonstrate that the TGFβ growth inhibition pathway is significantly hampered in IGHV3-21 CLL.

Genes with up-regulated expression in IGHV3-21 CLL cells according to GEP and RTQ-PCR validation experiments included VIPR1. Expression of VIPRs has been well documented in cells belonging to the immune system, where it promotes cell survival through a block of Fas/FasL-induced apoptosis. Of note,
VIPR1 belongs to the family of the G protein–coupled receptors, and its triggering by the specific ligand activates proteins of the Rho guanosine triphosphatase (Rho GTPase) family to transmit downstream signals. In this regard, the gene encoding for Rho family GTPase1 protein (RND1) was up-regulated in IGHV3-21 CLL, as it was another gene from the same family, the guanidine nucleotide exchange factor 1 (GEFI), the latter according to GEP data reported by Falt et al.

**GEP of homHCDR3–IGHV3-21 CLL**

A supervised analysis aimed at identifying the differential GEP of homHCDR3–IGHV3-21 versus nonhomHCDR3–IGHV3-21 CLLs was performed by applying a standard SAM statistical analysis. By setting the threshold of significance for the adjusted P value at 10⁻³ as made for multi-SAM, no differentially expressed genes were selected (D.M., personal communication, May 2006), thus suggesting that homHCDR3–IGHV3-21 and nonhomHCDR3–IGHV3-21 CLLs are more closely related than IGHV3-21 and non–IGHV3-21 cases. However, by lowering the stringency of the analysis and setting the statistical significance at an adjusted P value of less than 10⁻², we identified as few as 20 genes that differentiated homHCDR3–IGHV3-21 from nonhomHCDR3–IGHV3-21 CLL by cluster analysis. Among the genes of known function differentially expressed with the highest score, WNT-16 was found to be up-regulated in homHCDR3–IGHV3-21 CLL, as also validated by RTQ-PCR.

Members of the WNT gene family encode for secreted glycoproteins known to regulate normal B-cell differentiation and survival and have been found to be active and overall up-regulated in CLL. In particular, WNT-3, WNT-6, WNT-10, and WNT-16 were comparably expressed in UM and M CLls, while WNT-5b and WNT-14 levels were higher in UM CLL cells. The up-regulated expression of WNT-16 in homHCDR3–IGHV3-21 was paralleled by that of WNT-3, a gene selected as up-regulated in IGHV3-21 CLL cells by multi-SAM, although not at a ranking as high as to be included in the 122 best-correlated probes. Because an uncontrolled WNT signaling contributes to apoptosis impairment both in CLL and acute leukemias, the up-regulation of selected WNT members in IGHV3-21 and homHCDR3–IGHV3-21 CLL provides novel molecular evidence for the worse prognosis of this disease subset.

**References**

15. Matthews C, Catherwood M, Morris TC, Alexander HD. Routine analysis of IgVH mutational

**Acknowledgments**

Supported in part by Ministero della Salute (Ricerca Finalizzata IRCCS and Alleanza Contro il Cancro), Rome; Programmi di Ricerca di Interesse Nazionale (PRIN) and Fondo per gli Investimenti per la Ricerca di Base (FIRB), Ministero dell’Università e della Ricerca (MUR), Rome; Associazione Italiana contro le Leucemie linfomi e mieloma (AIL)-Venezia, Pramagguire group (M.D.B., fellowship); Novara-AIL Onlus, Novana; Ricerca Scientifica Applicata Regione Piemonte, Torino; Associazione Italiana per la Ricerca sul Cancro (AIRC), Milan; Piano di Ateneo per la Ricerca (PAR) 2005, University of Siena, Italy; The Leukemia & Lymphoma Society (White Plains, NY); Hairy Cell Leukemia Research Foundation (Shaumburg, IL); European Hematology Association 2004-Clinical Research Grant; and The Swedish Cancer Society.

**Authorship**

Contribution: R.B., M.D.B., D.B., and M.D. performed part of the IGHV gene mutation and gene expression profiling studies and contributed to data analyses and writing of the manuscript; D.C., F.F., D.G.E., R. Maffei, E.M.G., and M.T. performed part of the IGHV gene mutation studies and contributed to writing of the manuscript; A.Z. performed immunophenotypical studies and contributed to writing of the manuscript; R.R., A.G., R.F., R. Marasca, and G.G. provided part of the IGHV gene sequences and contributed to writing of the manuscript; G.D.P., M.I.D.P., P.B., and L.L. provided clinical data of patients and some of the CLL cases; D.M., G.P., and R.C. performed all the bioinformatic analyses of gene expression profiling; and V.G. coordinated the study and data analyses and wrote the manuscript.

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

Correspondence: Valter Gattei, Clinical and Experimental Hematology Research Unit, Centro di Riferimento Oncologico, I.R.C.C.S., Via Pedemontana Occidentale, 12, Aviano (PN), Italy; e-mail: vgatei@cro.it.