Selective influences in the expressed immunoglobulin heavy and light chain gene repertoire in hairy cell leukemia

Francesco Forconi,1 Elisa Sozzi,1 Davide Rossi,2 Surinder S. Sahota,3 Teresa Amato,4 Donatella Raspadori,1 Livio Trentin,5 Lorenzo Leoncini,2 Gianluca Gaidano,2 and Francesco Lauria1

1Sezione Ematologia e Trapianti, Università di Siena; 2Divisione di Ematologia, Dipartimento di Medicina Clinica e Sperimentale and Centro Interdisciplinare di Biotecnologie per la Ricerca Medica Applicata, Università del Piemonte Orientale Amedeo Avogadro, Novara, Italy; 3Cancer Sciences Division, University of Southampton, UK; 4Sezione di Anatomi PATologica, Università di Siena, and 5Dipartimento di Medicina Clinica e Sperimentale, Divisione di Ematologia-Immunologia, Università di Padova, Italy

Funding: the work was supported by an EHA clinical research grant 2004, and grants from the Hairy Cell Leukemia Foundation (Illinois, USA), Piano di Ateneo per la Ricerca (University of Siena), PRIN-MUR 2006 (Progetto di Ricerca di Interesse Nazionale - Ministry of University and Research), Ricerca Sanitaria Finalizzata, Regione Piemonte, Novara-AIL Onlus, Associazione Italiana per la Ricerca sul Cancro (AIRC, Italy), Siena-AIL Onlus. SSS is supported by the Leukaemia Research Fund UK.


Correspondence: Francesco Forconi, MD, DM, PhD, Sezione di Ematologia e Trapianti, Dipartimento di Medicina Clinica e Scienze Immunologiche, Università di Siena, AOUS, viale Bracci, 53100 Siena, Italy E-mail: forconi@unisi.it

The online version of this article contains a supplementary appendix.

ABSTRACT

Background
We previously reported ongoing mutational and isotype switch events in the immunoglobulin (Ig) heavy chain (H) locus in hairy cell leukemia. Those analyses raised questions on the incidence and type of selective influences occurring on the tumor B-cell receptor of hairy cell leukemia.

Design and Methods
To further investigate this issue, we examined the full IGH and κ and λ light chains (IGκ and IGL) variable and constant region transcripts expressed in a large cohort of patients with hairy cell leukemia (n=88).

Results
Multiple IGH isotypes were expressed in 46/56 (82%) cases of hairy cell leukemia. Comparison of tumor with normal B-cell repertoires revealed preferential usage of IGHV3-21, IGHV3-30 and IGHV3-33 in hairy cell leukemia (p=0.001, p=0.003 and p=0.001, respectively). Light chain analysis demonstrated preferential IGL use with an inverted IGHκ:IGL ratio (0.7:1) and universal usage of IGLJ3. Analysis of LCDR3 junctions revealed highly homologous motifs in 40% of IGL.

Parallel analysis of IGH and IGL showed selective pairing of IGHV3-21/30/33 segments to specific LCDR3-J3 subsets (p=0.008). Of 40 cases of hairy cell leukemia, 38 had mutated IGH and/or IGL, with variations in 13/13 cloned cases, while two had 100% unmutated IGH and IGL.

Conclusions
Overall, biased IGH usage, preference for IGL with universal IGLJ3 usage and a high incidence of LCDR3 homologous motifs suggest selective influences on the B-cell receptor of hairy cell leukemia. Ongoing mutations and isotype switching suggest that influences occur on the tumor B-cell receptor at ectopic sites.

Key words: hairy cell leukemia, immunoglobulin, IGH, IGL, IGκ, heavy chain, light chain.


©2008 Ferrata Storti Foundation. This is an open-access paper.
Introduction

Hairy cell leukemia (HCL) is a rare, chronic B-cell neoplasm characterized by leukemic hairy cells present in blood, bone marrow, and splenic red pulp, with atrophy of white pulp. Lymph node involvement is infrequent. HCL is typically associated with markers of activation, which include expression of CD25, CD11c, FMC7, and CD103 at high intensity. A distinctive feature of HCL is expression of multiple surface (s) immunoglobulin (Ig) isotypes, although their prevalence in HCL has not been fully mapped. 1 2

B-cell tumors preserve the B-cell receptor (BCR) features of the originally transformed cell. 1 Consequently, immunoglobulin gene (IG) analysis delineates the critical events of clonal development and defines the IG heavy (H) and light (k) repertoire selected by specific tumor entities. 3 In some instances, IG analysis may also have prognostic value, 3,4 and the selected IGH/IGL or IGH/IGK pairs can associate with specific BCR structure and clinical behavior. 5 Selective stimuli to the tumor BCR may be of different types, including viral or bacterial antigens, or, in germinal center-derived lymphomas, stromal elements acting on N-glycosylated residues acquired by somatic mutation. 6 Analysis of the selective influences on the tumor BCR has often been hampered in HCL by the rarity of the disease, and only small series of cases have been analyzed to date. 1

In small series of HCL, we and others have observed that most HCL carry mutated IGH variable region (V) genes, with low levels of intraclonal heterogeneity. Only a minor subset of HCL has unmutated IGHV genes. 2,5,9 Both mutated and unmutated subsets of HCL express multiple slgH isotypes with no evidence of subpopulations. 2 Also, activation-induced cytidine-deaminase is expressed in HCL, and Igg sterile transcripts are produced prior to class switch deletional recombination. 7 However, hairy cells fail to express the germinal center markers CD38, CD10 and BCL6, or the M13 forward and reverse primers were used to

Design and Methods

Patients

Peripheral blood samples were collected from 88 patients with HCL. In all instances, the specimens were collected at diagnosis before specific therapy. The diagnosis of typical HCL and variant HCL was based on peripheral blood morphology, flow cytometry and bone marrow immunohistochemistry according to the World Health Organization classification. 12 Informed consent was obtained in all cases and the study was approved by the local Institutional Review Board.

Phenotypic analysis

Peripheral blood mononuclear cells were obtained by Lymphoprep (Nycomed Pharma, Oslo, Norway) gradient separation. Immunophenotypic studies were carried out on these cells by direct immunofluorescence techniques with a large panel of antibodies. 7 Expression of CD27, CD38 and slgH isotypes on HCL was determined by three-color staining with F(ab)’2 anti-slgG, anti-slgM, anti-slgD, and anti-slgA antibodies. 7 Expression of slgk was determined by three color staining with fluorescein isothiocyanate (FITC)– conjugated F(ab)’2 anti-slgk, phycoerythrin (PE)-conjugated F(ab)’2 anti-slgk and peridinin-chlorophyll protein (PerCP)-conjugated anti-CD20. Previously described procedures were used to avoid non-specific binding. 7 Data were acquired and analyzed and antigen expression was defined as described elsewhere. 7

Polymerase chain reaction amplification of IGHVDJ, IGKV and IGLV transcripts and sequence analysis

Total RNA was isolated from peripheral blood mononuclear cells and cDNA prepared as described previously. 7 The full tumor IGHVDJ transcripts were amplified by PCR with a mixture of Leader-VH-mix primers and a constant-region primer. 7 The full tumor IGKVDJ or IGLVJ transcripts were identified by PCR using a mixture of 5’ primers specific for known IGK or IGL leader sequences (IGKV- or IGLV-leader mix), together with a 3’ primer specific for either the IGK or IGL constant-region. 10 Amplified products were run on agarose gel and purified with a JetQuick GEL extraction kit (Biozym, Hagele, Germany). The tumor IGHVDJ, IGKVJ and IGLVJ sequences were identified by direct sequencing and/or after cloning the purified band. Ligation and cloning were performed with pGEM®-T Easy Vector System II and JM109 competent cells (Promega, Milan, Italy). The tumor IGHVDJ, IGKVJ and IGLVJ sequences were identified by direct sequencing and/or after cloning the purified band. Ligation and cloning were performed with pGEM®-T Easy Vector System II and JM109 competent cells (Promega, Milan, Italy). Sequencing was carried out using the v1.1 Big Dye Terminator Ready Reaction sequencing kit (AB Applied Biosystems, Applera Italia, Monza, Italy), on an ABI Prism 310 genetic analyzer (PerkinElmer, Warrington, UK). Direct sequencing was performed with the 3’ primer on the constant region and the identified sequence was confirmed with the family specific leader 5’ primer, which also identified identification of the full V-gene transcript. When cloning was performed, M13 forward and reverse primers were used to
sequence in both directions. The data were analyzed using Chromas 1.51 software and aligned to the 2005 updated V-Base and ImMunoGeneTics (IMGT) databases.6,15

IGHV, IGKV and IGLV gene usage and mutation patterns were analyzed as previously described.7 IMGT nomenclature was used to assign IG gene use,14 since this allowed comparison with data from previously used nomenclatures for IGH, IGK and IGL genes.16,17 Lengths in the IGH and IGK/LCDR3 (HCDR3, KCDR3 and LCDDR3, respectively) were calculated according to IMGT criteria.14 IMGT criteria and nomenclature were also used for IGHD determination, again allowing comparison to segments with other designations from the literature.14 N-addition of G and C at the joining ends of the V(D)J junction ([Ngc] Online Supplementary Table S1) was performed to investigate TdT activity. Recurrent amino acid sequence motifs in HCDR3, KCDR3 and LCDR3 were sought using the ClustalW tool (at http://www.ebi.ac.uk/clustalw/). Amino acid identity >60% in the HCDR3 or >80% in the K/LCDR3 was required for the inclusion of an IGH or an IGL/L chain in the same subset, respectively.20 Intraclonal heterogeneity was asessed in the cloned products and was distinguished from Taq infidelity by an increased frequency compared to Taq error rate and by the finding of the same mutation in more than one clone.1 If only direct sequencing was performed, the tumor IGHV, IGKV and IGLV sequences were confirmed by replicate RT-PCR and sequencing. The incidence of potential novel N-glycosylation sites in IGHVDJ, IGKVJ and IGLVJ transcript sequences was assessed as previously described.6,19 Statistical analyses were performed using \( \chi^2 \) or Fisher’s exact tests. \( p \) values <0.05 were considered statistically significant.

**Results**

**Analysis of expressed IGHV region transcripts and sIgH isotype proteins**

The expressed tumor IGHV sequences were identified in 83/88 HCL (Online Supplementary Table S1). The distribution of IGHV families was similar to that of normal B-cells, but use of individual IGHV gene segments showed differences from that of the normal B-cell repertoire (Figure 1A). The IGHV gene segments most frequently used in HCL were IGHV3-30 (13/83, 16%), IGHV3-33 (9/83, 11%), IGHV3-23 (7/83, 8.5%), IGHV3-21 (7/83, 8.5%), IGHV4-30/4-31 (5/83, 6%), IGHV3-34 (5/83, 6%). Among these segments, the usage of IGHV3-21, IGHV3-30 and IGHV3-33 was significantly increased compared to the normal B-cell repertoire \( (p<0.001, p=0.005 \text{ and } p=0.001, \text{ respectively}) \).6 Other IGHV segments were used less frequently in HCL than in normal B-cells (Figure 1A), although the differences were not statistically significant, for any of the segments individually, likely due to the number of cases investigated. However, by combining results of this study with those of all published HCL IGHV gene sequences \( (n=164, \text{ Online Supplementary Figure S1A}) \), we confirmed the overuse of IGHV3-21, IGHV3-30 and IGHV3-33, and demonstrated the significantly reduced use of IGHV1-18 \( (p=0.02) \) and IGHV3-53 \( (p=0.02) \).6,19,21,22

Phenotypic analysis of surface IgM, IgD, IgG and IgA was performed in 56 HCL (Online Supplementary Table S1). The majority (46/56, 82%) co-expressed multiple pre- (IgM±D) and post-switch (IgG±A) isotypes, indicating that multiple isotype expression is a dominant feature in HCL.6 Of 63 HCL tested, CD38 was negative in all cases, and the memory B-cell marker CD27 was negative in 58/63 cases. Interestingly the five CD27-positive cases (cases HCL8, HCL17, HCL72, HCL82, and HCL87) differed from all the remaining HCL investigated, as they had features consistent with the diagnosis of variant HCL, including high white blood cell counts with prolymphocytic morphology and CD25...
negativity of the peripheral blood hairy cells. These data confirm that lack of surface CD27 and CD58 is a feature of typical HCL.

**Analysis of HCDR3 junction**

Sixty-nine cases of HCL were evaluable for the HCDR3 junction (Online Supplementary Table S1 and Figure 2). IGHD gene segment analysis revealed significantly increased use of IGHD 1 (11/69 cases, 16%) and IGHd6 (12/69, 17%) families (p=0.001 and 0.02, respectively). Analysis of specific segments revealed additional notable biases. In fact, the IGHd1 family used almost exclusively IGHD1-26 (10/11 cases) and the IGHd6 family frequently utilized IGHd6-49 (5/12 cases). This observation is remarkable, since neither of these segments has been reported to be used in normal B cells. In addition, significantly increased selection of IGHd3-9 (4/69, p=0.004), IGHd3-10 (7/69, p=0.0005) and IGHd4-17 (5/69, p=0.0009) was also documented. Again, IGHD3-3, IGHD3-9, and IGHd4-17 are not reportedly selected in the functional normal B-cell repertoire, and only IGHD3-10 is restricted to 0.5% of all D segments.

IGHJ gene segments used in the expressed HCL repertoire distributed similarly to those in the normal B-cell repertoire (Online Supplementary Table S1). However, the number of sequences available in the literature (n=27) precludes any statistical analysis. Among functional B-cells from published data, dark gray columns: HCL; light gray columns: normal B-cell repertoire. Asterisks indicate regions with identified significant differences between HCL and the normal B-cell repertoire. The D1 family was represented by D1-26 in 10/11 cases. The D6 family was represented by D6-19 in 5/12 cases. D1-26 and D6-19 are not reportedly used by normal B cells.

**Analysis of sIgk and sIgl protein expression and IGKV and IGLV region transcripts**

Seventy of 83 HCL were characterized for sIgk and sIgl protein expression (Online Supplementary Table S1). Expression of sIgk was observed in 41/70 (59%) HCL with an Igk/Igl ratio of 0.7:1, indicating preferential secondary rearrangement of Igk. The expressed IGK and IGL tumor transcript sequences were investigated in 45/70 cases (25 HCL expressing sIgk and 22 HCL expressing sIgl). In two additional samples (HCL30 and HCL38), IGL transcripts from the non-functional allele were amplified and sequenced. Four HCL (HCL4, HCL42, HCL70, and HCL83) co-expressing double IGL or double IGK, or IGL/IGK, were excluded from the analysis. Among IGK gene segments (Online Supplementary Table S2 and Figure 1B), IGKVD3-39 (01/2/02) was the one most frequently used (6/23, 26%), and its usage was significantly higher than in normal B-cells (p=0.03). The distal gene segments IGKVD1-17 (case HCL63) and IGKVD6-21 (case HCL11), which were not reported in the normal B-cell repertoire, were both used once. The other IGK segments distributed similarly to normal B-cells.

Among functional IGLV gene segments, IGLV2-14 was most frequently used (4/22, 18%), followed by IGLV1-47 (3/22, 14%), IGLV1-40, IGLV1-44, IGLV1-51, IGLV2-11 and IGLV3-21 (2/22, 9% each) (Online Supplementary Table S2 and Figure 1C), with an overall distribution similar to that of the normal B-cell repertoire.

**Analysis of KCDR3 and LCDR3 junctions**

Twenty-one HCL were evaluable for the KCDR3 junction. IGKJ genes were used by HCL in a fashion similar to that in the normal B-cell repertoire (Online Supplementary Table S2). KCDR3 length ranged from 7 to 11 amino acids (median 9) and 11/21 KCDR3 had identical pl (range 6.5-13, median 13). Overall analysis of KCDR3 amino acid sequences reflected pl similarities and identified three subsets, all with 88.8% sequence identity (Online Supplementary Table S2). Subset 1K (HCL28 and HCL2) harbored IGKV1D-33/17 rearrangements (QYQNLP[L/R]TG), which associated with IGK/V3-23 (HCL28) or IGK/V1-02 (HCL2). Subset 2K (HCL7/330 and HCL67) harbored IGKV3-20/J1-2/1 rearrangements (QYGRSP[Q/Y]YT), which associated with IGK/V2-05 (HCL7/330) or IGKV4-34 (HCL67). Subset 3K (HCL19 and HCL68) harbored IGKV1-17/1D-17/J1-2/1 rearrangements (QHNSYP[Q/Y]YT), which associated with IGK/V3-24 (HCL19) or IGKV4-34 (HCL68).

Twenty-two HCL were evaluable for the functional LCDR3 junction. Remarkably, among IGLJ segments, virtually all cases used IGLJ3 (21/22, 95.5%) (Online Supplementary Table S2 and Figure 3). The universal use of IGLJ3 was higher than expected by chance alone (50%), and significantly higher than its frequency (34%) in normal B cells (p=0.00000001). Remarkably, 2/2 HCL with non-functional LCDR3 junctions (HCL30 and HCL33) failed to use IGLJ3, further indicating strikingly selective influences on the functional repertoire. In particular, HCL30 used IGLJ1 with a TAG stop at joining codon 115, whereas HCL33 used IGLJ6 with and out-of-frame rearrangement in codon 115 (Online Supplementary Table S2). LCDR3 length ranged from 9 to 13 amino acids (median 11) and the pl ranged from 4.4 to 13 (median 13), with 13/22 LCDR3 having an
identical pl.

Analysis of LCDR3 amino acid sequences identified three HCL subsets (Figure 4). Subset 1L (HCL35, HCL22, HCL44, HCL49, and HCL40) harbored IGLV4-44/47-J3 rearrangements with ≥85% homologous LCDR3 (Figure 4). Of note, these HCL paired only with the highly identical IGHV3-30 or IGHV3-33 segments in four out of five cases, the exception being case HCL22, which used IGHV3-24. Remarkably, the selective clustering of IGHV3-21/30/33 segments (5/5, 100%) within this subset was significant, when compared to pairing of IGHV3-21/30/33 segments with the remaining CDR3 investigated for this subset was significant, when compared to pairing of IGHV3-21/30/33 segments with the remaining CDR3 investigated for IGLV-J (8/17, 47%, p=0.05) or the total IGLV/J (13/38, 34%, p=0.005) rearrangements (Online Supplementary Table S2). Subset 2L (HCL27 and HCL32) harbored IGLV3-21-J3 rearrangements with 91% homologous LCDR3 (QVVDSSS-DH[W/V/V]), and associated with either IGHV3-30 or IGHV3-23. Subset 3L (HCL64 and HCL23) harbored IGLV4-40-J3 rearrangements with 81.8% homologous LCDR3 (QSYD[S/N]SL[SG/TR]SGV), and associated with IGHV4-34 or IGHV3-24. Overall, the overused IGHV3-60, IGHV3-33 and IGHV3-21 genes clustered within the identified LCDR3 sets (7/9 cases, 78%), and not with the CDR3 from the remaining IGLV/J rearrangements investigated (9/34, 26%, p=0.008).

Together with Igα light chain preference and universal IGLJ3 use, these data further suggest that selection events are dominant in the Igα light chain of HCL. TdT activity was absent in almost 50% light chain rearrangements (20/41 light chain junctions). Since TdT activity decreases during B-cell ontogeny, lack of N-addition/deletion of G and C at the ends of V-J junction sequences is suggestive of receptor revision in the light chain, after downregulation of the TdT enzyme.41

**Somatic mutation analysis of IGHV, IGKV and IGLV**

The mutation status of IGHV tumor transcripts was evaluated in 83 HCL (Online Supplementary Table S1). Overall, IGHV genes carried variable tiers of mutations (range 80.95-100%; median 96.56%, mean 95.97%). Three of 83 HCL had completely unmutated IGHV genes (100% homology to germline). However, by using the arbitrary 2% cut-off that defines mutational status in other lymphoproliferative disorders and that is being used for clinical correlates in ongoing clinical studies,26 the unmutated HCL subgroup expanded to 17/83 HCL (20.5%).

The distribution of individual IGHV genes varied among mutated and unmutated cases of HCL (Online Supplementary Table S2 and Figure 5A). In particular, IGHV3-30, the segment most frequently used in HCL, showed a selective tendency to distribute in unmutated rather than mutated HCL cases (6/17, 35% unmutated HCL vs. 7/66, 10% mutated HCL, p=0.01). IGHV4-34 was also more frequently used in unmutated HCL (3/17, 18%) than in mutated cases (2/66, 3%) (p=0.02). IGKV mutation status (n=23) (Online Supplementary Table S2 and Figure 5B) was variable (homology range 93.61%-100%, median 97.26%, mean 97.31%), and 2/23 cases had completely unmutated IGKV genes. Using the arbitrary 2% cut-off value, 9/23 (39.1%) IGKV were considered unmutated. IGLV mutation status was evaluable in 22 cases with functional IGLV rearrangements (Online Supplementary Table S2 and Figure 5C). IGLV homology to germline ranged from 89.9% to 100% (median 97.12%, mean 96.83%), and 1/22 HCL had completely unmutated IGLV genes. Using the arbitrary 2% cut-off, 8/22 (36.4%) IGLV were unmutated.

Parallel assessment of IGHV and IGKV/IGLV mutation status was feasible in 40 cases of HCL (Online Supplementary Table S2). The IGHV-IGKV/IGLV mutation status was concordant in 22/40 HCL. Eighteen cases were mutated in both IGHV and IGKV/IGLV.

---

**Figure 3. IGLJ segment usage in Igκ HCL.** The incidence of IGLJ segments in the functional IGLV transcripts expressed by HCL was compared to the functional repertoire of normal B cells from published data.43 Dark gray columns: HCL; light gray columns: normal B-cell repertoire. *IGHJ3 segment use in HCL was universal (≥95% cases) and significantly superior to IGLJ3 use in normal B cells.

**Figure 4. Subsets of highly similar LCDR3 in HCL.** Amino acid sequences are shown for all HCL cases within a LCDR3 subset. The rearranged IGLV and IGLJ and the associated LCDR3 segments with the remaining IGLJ3 segments. (*, **, ***): *selective distribution of IGLV-J subsets 1L and 2L compared to HCL with light chain CDR3 not belonging to subset 1L; **selective distribution of IGLV-J subsets 1L and 3L compared to HCL with light chain CDR3 not belonging to subset 1L; ***selective distribution of IGHV3-21, 3-30 or 3-33 in HCL from subset 1L compared to IGLJ-positive HCL not belonging to subset 1L; **selective distribution of IGHV3-21, 3-30 or 3-33 in HCL from subset 1L compared to IGLJ-positive or IGLJ-negative HCL with light chain CDR3 not belonging to subset 1L; **selective distribution of IGHV3-21, 3-30 or 3-33 in HCL from subset 1L compared to IGLJ-positive or IGLJ-negative HCL with light chain CDR3 not belonging to these subsets.

---

The incidence of IGLV-J subsets in HCL (Online Supplementary Table S2). The use of the arbitrary 2% cut-off that defines mutational status in other lymphoproliferative disorders and that is being used for clinical correlates in ongoing clinical studies,26 the unmutated HCL subgroup expanded to 17/83 HCL (20.5%). The distribution of individual IGHV genes varied among mutated and unmutated cases of HCL (Online Supplementary Table S2 and Figure 5A). In particular, IGHV3-30, the segment most frequently used in HCL, showed a selective tendency to distribute in unmutated rather than mutated HCL cases (6/17, 35% unmutated HCL vs. 7/66, 10% mutated HCL, p=0.01). IGHV4-34 was also more frequently used in unmutated HCL (3/17, 18%) than in mutated cases (2/66, 3%) (p=0.02). IGKV mutation status (n=23) (Online Supplementary Table S2 and Figure 5B) was variable (homology range 93.61%-100%, median 97.26%, mean 97.31%), and 2/23 cases had completely unmutated IGKV genes. Using the arbitrary 2% cut-off value, 9/23 (39.1%) IGKV were considered unmutated. IGLV mutation status was evaluable in 22 cases with functional IGLV rearrangements (Online Supplementary Table S2 and Figure 5C). IGLV homology to germline ranged from 89.9% to 100% (median 97.12%, mean 96.83%), and 1/22 HCL had completely unmutated IGLV genes. Using the arbitrary 2% cut-off, 8/22 (36.4%) IGLV were unmutated.

Parallel assessment of IGHV and IGKV/IGLV mutation status was feasible in 40 cases of HCL (Online Supplementary Table S2). The IGHV-IGKV/IGLV mutation status was concordant in 22/40 HCL. Eighteen cases were mutated in both IGHV and IGKV/IGLV.
Four were unmutated in both IGHV and IGKV/IGLV. Of these, two cases displayed both IGHV and IGKV/IGLV rearrangements with a 100% homology to closest germline, confirming the existence of a very minor subset of HCL with completely unmutated IG genes.

The IGHV-IGKV/IGLV mutation status appeared discordant in 18 cases of HCL (12 IGHV mutated-IGKV/IGLV unmutated cases and six IGHV unmutated-IGKV/IGLV mutated cases). However, 16/18 discordant cases carried some level of mutation (homology to germline <100%) in both IGHV and IGKV/IGLV, suggesting that the variations likely represent true mutations and not polymorphisms. Only two discordant cases (HCLS7 and HCL81) carried heavily mutated IGHV (94.79% homology) or IGKV (94.98%) genes coupled to completely unmutated (100%) IGKV or IGHV genes, respectively. These two particular cases might be representative of an antigen experienced BCR (with mutated IGH or IGK) rescued after secondary recombination of the IGK or IGH chain on the second allele, and suppression of the first functionally rearranged allele (receptor revision). 30,31

Analysis of intraclonal heterogeneity of IGHV, IGKV and IGLV

Cloning and sequencing of IGHV and IGKV or IGLV transcripts was performed in 12 cases of HCL (Table 1). In all cases, cloning confirmed the results of direct sequencing. Cloning of IGHV transcripts revealed intraclonal variations within the same or different tumor isotypes in 11/12 cases, including 2/2 cases with >98% and <100% IGHV homology to germline (cases HCL7 and HCL35). Cloning of the paired IGK/L tumor sequence confirmed intraclonal variations in the light chain of the same 11/12 cases. Using stringent criteria for ongoing activity (i.e. ≥2 identical variations repeated in separate clones), intraclonal heterogeneity was restricted to 3/11 cases. Lack of repetitions in the remaining 8/11 cases may be due to the mutational frequencies in the light chain genes, which, in normal individuals, are generally lower than in the IGHV genes. 32 Only HCL38, having both IGHV and IGLV genes with 100% homology to germline, did not display intraclonal mutations in either IGHV or IGLV genes.

Discussion

Our immunogenetic analysis of the expressed IGH and IGK/L repertoire in a large cohort of patients has implications for the origin of HCL, and indicates that IG selection may play an important role in the pathogenesis of this disease. IGHV analysis shows that HCL is characterized by biased usage of IGHV3-21, IGHV3-30 and IGHV3-33, in agreement with prior analyses on smaller HCL groups. 33 IGHV3-30 and IGHV3-33 are highly homologous segments with a single amino acid difference in the CDR2, are recognized by the same anti-CDR1 (B6) monoclonal antibody, and thus may share the ability to bind identical antigens. 34 Indeed, IGKV family members, including IGHV3-30 and IGHV3-33, react with common bacterial superantigens, such as modified staphylococcal protein A, 34 or with the natural staphylococcal enterotoxin A, which is sufficient to induce survival of IGHV3-expressing B cells by low-affinity binding. 35 Furthermore, IGHV3-30 is reportedly involved in the immune response against Toxoplasma infection, providing additional clues to potential antigens sustaining HCL. 36 Clearly, the selective influences active in HCL appear to follow routes that are different from those of other B-cell neoplasms, in particular from the extensively investigated IgM-positive chronic lymphocytic leukemia (CLL). 19 For example, IGHV1-69 was totally absent in HCL from
Table 1. Intraclonal variability in hairy cell leukemia.

<table>
<thead>
<tr>
<th>Code*</th>
<th>IGHV</th>
<th>Homology (%)</th>
<th>Intraclonal heterogeneity in IGHV</th>
<th>IGK/LV</th>
<th>Homology (%)</th>
<th>Tumor/total IGL</th>
<th>Intraclonal heterogeneity in IGK/LV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCL9</td>
<td>3-07</td>
<td>92.45</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL25/206</td>
<td>3-23</td>
<td>91.31</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL7/330</td>
<td>2-05</td>
<td>98.96</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL35</td>
<td>3-30</td>
<td>98.49</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL44/42</td>
<td>3-33</td>
<td>93.50</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL40</td>
<td>3-30</td>
<td>97.61</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL38/283</td>
<td>3-30</td>
<td>100.00</td>
<td>-</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL3/266</td>
<td>1-02</td>
<td>96.18</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL38/283</td>
<td>3-30</td>
<td>98.53</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
<tr>
<td>HCL26/216</td>
<td>3-23</td>
<td>91.66</td>
<td>+</td>
<td>KVI-039(012/02)</td>
<td>97.49</td>
<td>11/11</td>
<td>3</td>
</tr>
</tbody>
</table>

Intraclonal heterogeneity in the cloned IGHV products was defined positive when the same mutation was present in more than one clone from different isotype transcripts.2,7,29 The HCL cases with previously published details on intraclonal variation in tumor IGH isotypes are listed with the original code after the slash.2,7,29

our and previously published series (total 164 IGHV sequences),2,7,8,21,22,37 while it dominates the unmutated subset in CLL (Online Supplementary Figures S1A and S1B).19 Similarly, IGHV4-34 is used predominantly in a mutated conformation by CLL, but was preferentially unmutated in the HCL cases in our series (Figure 5A), and even more significantly when all published HCL sequence were pooled (p=0.0002).2,7,8,21,22,37 Lack of apparent HCDR3 stereotypy in HCL is another feature of variance with CLL, which, conversely, frequently groups cases with shared HCDR3 structural features.4,18,38 This also indicates that antigen drive may not rely on HCDR3-mediated interactions in HCL. Conversely, the low mutational rate of IGHV genes expressed in HCL, particularly of IGHV3-30 and IGHV3-33 segments, which represent almost 50% of all unmutated HCL, suggests that selective influences may be related to the IGHV segment itself.34,35

Analysis of Ig light chains provides novel evidence that HCL is characterized by selection events in the tumor BCR. In fact, HCL display: (i) an inverted Igκ: Igλ ratio (0.7:1); (ii) universal usage of IGLJ3 in the functional slgκ expressors; and (iii) subsets with highly homologous KCDR3 and LCDR3. The preferential usage of Igκ light chain in our large series is consistent with prior independent studies,19,28 and can be considered a unique feature of HCL. In the normal B-cell repertoire and in other B-cell neoplasms, Igκ is the most frequently used light chain.41,42 On these bases, our results suggest that HCL requires selective usage of Igκ. The functional implications of Igκ selection in HCL remain speculative. The observations that (i) almost 50% HCL expressing slgκ utilize IGHV3-21, IGHV3-30 or IGHV5-35 only; (ii) virtually all HCL expressing Igκ utilize IGLJ3; and (iii) 40% HCL expressing Igκ display LCDR3 sets with shared structural features, suggest that HCL expressing Igκ may recognize common antigens requiring homologous LCDR3-J3 stretches.

LCDR3 identical motifs were documented within IGHV3-21, IGHV3-30 or IGHV3-21 HCL (Table 1). In public databases of normal or autoreactive B-cells, we identified motifs identical to LCDR3 from HCL sets 1L or 2L, although specific antigen reactivities were found only in set 1L.2 The molecular triggers of the Igκ bias and LCDR3-J3 selection in HCL are unknown. One possibility of Igκ selection is that the Igκ-to-Igλ shift may derive from secondary rearrangements with rescue of a new light chain in the periphery. Secondary rearrangements of Igλ after Igκ deletion have been observed in cases of Igλ-positive B-cell neoplasms.45-47,49

The observation of absent Nqc incorporations due to absent TdT activity in almost 50% HCL light chain rearrangements favors the hypothesis that secondary rearrangements occur in the periphery.27 We have observed the potential ability of HCL to revise light chains in cases co-expressing mutated double light chain transcripts and/or protein with peripheral up-regulation of RAG-1.53 Also, double productive and functional Ig light chain expression has been described in one independent case of HCL and is putatively the consequence of peripheral receptor revision.23

Classically, mutational status of IGHV genes distinguishes whether the B-cell has encountered antigen in the germinal center with a T-cell-dependent reaction or whether it derives from antigen-naïve B cells.7 However, there is evidence that the BCR can also interact with antigen in a T-independent pathway and accumulate a low level of somatic mutations ectopically, likely in the marginal zones.21,22 Several observations from both Ig heavy and light chain rearrangements indicate that HCL cells have experienced antigen stimulation. First, the majority of HCL are characterized by somatic mutations. Second, cloning of IGV region transcripts confirms the existence of low levels of intraclonal heterogeneity also in cases with 98% ≤ homology <100% to germline.19 Third, the vast majority (82%) of HCL co-express multiple pre-switch (IgM±D) and post-switch (IgG±A) slgH, independently of mutational status and while activation-induced cytotoxic deaminase is expressed.7 Histological findings and lack of CD27 and CD38 suggest that, in typical HCL, the mutational and switch events are unlikely to occur in the germinal cen-
In conclusion, the present largest immunogenetic survey identifies the Ig repertoire selected in HCL. The remarkable preference for Iga with universal IGLJ3 use and a high incidence of LCDR3 homologous motifs further clarifies the selective influences present in the BCR of HCL. Whether this distribution of repertoire in HCL reflects that of a postulated normal B-cell counterpart remains to be clarified, and may follow from a more accurate characterization of normal B-cell subsets.

**Authorship and Disclosures**

FF designed the study, collected and processed the samples, performed experiments, analysed results and wrote the manuscript; ES, DR, SSS, TA, DR, FT and LL collected and processed the samples, performed experiments, and analyzed results; LT provided tumor samples, performed experiments and analyzed data; GG designed the study, provided tumor samples, analyzed data and wrote the manuscript; FL provided tumor samples, supervised the project and manuscript writing. The authors reported no potential conflicts of interest.

**References**

15. MRC wss. VBASE database. Available at http://www.mrc-cpe.cam.ac.uk.
20. Brezinschek HP, Foster SJ, Brezinschek RJ, Domer T, Domiati-Saad R, Lipsky PE. Analysis of the human


