Somatic Diversification and Selection of Immunoglobulin Heavy and Light Chain Variable Region Genes in IgG^+CD5^+ Chronic Lymphocytic Leukemia B Cells

By Shiori Hashimoto,* Mariella Dono,* Mariko Wakai,* Steven L. Allen,* Stuart M. Lichtman,* Philip Schulman,* Vincent P. Vinciguerra,* Manlio Ferrarini,‡ Jack Silver,* and Nicholas Chiorazzi*

*From the *Departments of Medicine, North Shore University Hospital and Cornell University Medical College, Manhasset, New York 11030; and the ‡Laboratory of Clinical Immunology, Istituto Nazionale per la Ricerca sul Cancro, Genoa, Italy 16132*

Summary

Chronic lymphocytic leukemia (CLL) is characterized by the clonal expansion of CD5-expressing B lymphocytes. Most studies have found that these leukemic $CD5^+$ B cells, like their normal counterparts, use immunoglobulin (Ig) variable (V) region genes that exhibit minimal, if any, somatic diversity. These and other observations have suggested that $CD5^+$ B cells may be incapable of generating Ig V gene diversity, and therefore may not be able to develop higher affinity binding sites that could be selected by antigen. However, most of the studies of CLL and normal $CD5^+$ B cells have focused on IgM-producing cells. Since somatic mutations are most often seen in B cells that have undergone an isotype class switch, we analyzed the Ig heavy (H) and light (L) chain variable region genes of seven IgG^+CD5^+ CLL B cells to determine if somatic diversification and antigen selection had occurred. The data derived provide evidence for skewed use, somatic diversification, and antigenic selection of the Ig V region genes. Nonrandom use of both H and L chain V region genes was manifested by an overrepresentation of V_H4 and $V_\kappa I$ family genes and the underrepresentation of the J_H4 gene segment. Furthermore, V_H4 gene use was restricted to only two family members (4.21 and 4.18). In four of the seven cases, the V_H and V_L genes displayed $\geq$5% difference from the most homologous known germline counterparts. Polymerase chain reaction and Southern blot analyses performed in two of these patients demonstrated that their unique V_H CDR2 and adjacent sequences were not present in their germline DNA. In addition, a significant level of diversity was seen in the rearranged DJ_H segments and at the V_L-J_L junctions of every patient that occurred both at the time of recombination and subsequently. The localization of replacement changes to complementarity determining regions of some patients suggested that antigen selection had occurred. Furthermore, the mutations identified in the V_H and V_L genes of each individual patient were strikingly similar, both in number and location. Collectively, the data indicate that a subset of $CD5^+$ CLL B cells can display Ig V region gene mutations. In addition, they are consistent with the notions that in some cases antigen selection of these mutations may have occurred, and that antigen stimulation may be a promoting factor in the evolution of certain CLL clones.

Chronic lymphocytic leukemia (CLL)[1] is characterized by an overexpansion of a clone of human B cells that almost invariably expresses surface membrane CD5 (for a review see reference 1). Over the years, investigators have studied these cells to understand both the normal physiology of this subset of human B cells as well as the abnormal features of the B lymphocytes expanded in this disease. Studies to date suggest that the B cells that are clonally expanded in CLL use a biased set of Ig V genes (for a review see reference 2) to code for low affinity, polyreactive autoantibodies (3–5), predominantly of the IgM class. Although numerous studies have noted that these Ig V genes rarely undergo somatic mu-

[1] *Abbreviations used in this paper:* CDR, complementarity determining region; CLL, chronic lymphocytic leukemia; FR, framework region; RT, reverse transcriptase.

This work was presented in part at the International Conference on Polyreactive and Monoreactive Antibodies in Natural Immunity and Autoimmune Diseases held at the National Institutes of Health, Bethesda, MD, 23–25 September 1992.

tation (2), this principle was challenged by the study of Cai et al. (6), which documented V_H gene mutations among a cohort of CLL patients whose malignant clone used the V_H251 gene. Surprisingly, however, a follow-up study (7) of a larger, well-characterized group of patients whose leukemic cells used the same gene failed to confirm these findings, thus leaving this dilemma unresolved.

It is important to note, however, that most of these studies of $CD5^+$ CLL B cells have dealt with IgM-producing cells. Since antigen stimulation resulting in somatic diversification may occur more often at the time of or among isotype-switched B cells (8), $CD5^+$ B cells producing non-IgM antibodies may exhibit different features, both at the protein and nucleic acid levels. To test this hypothesis and to help resolve the issue of the capacity of these cells to somatically diversify their Ig V genes, we have analyzed the Ig V genes used by $CD5^+$ CLL cells that produce IgG antibodies. The data demonstrate an overrepresentation of Ig V_H and V_L family genes and an underrepresentation of a frequently employed J_H segment. In addition, there is strong evidence that the Ig V gene segments of both the H and L chains from several of these $CD5^+$ CLL clones have undergone somatic mutation. Finally, the location of replacement substitutions suggests that some of these mutations may have been antigen driven and selected.

Materials and Methods

CLL Cells and Heterohybridomas. Seven patients with CLL whose $CD5^+$ leukemic B cell clones produced IgG were studied. As recently reported (9), each patient exhibited clinical features typical for CLL, with stages ranging from 0 to IV (Rai classification; 10). In addition, all exhibited an expanded population of circulating IgG^+CD5^+ B cells as determined by immunofluorescence, although the percentages of IgG^+CD5^+ coexpressing cells varied among the various patients (range: 64–92%). Serologic and immunofluorescence studies demonstrated that five of these CLL clones produced IgG1 and two IgG3 (11).

Heterohybridomas made with these cells were screened for the secretion of Ig H and L chains that corresponded to the original $CD5^+$ CLL B cell, and appropriate lines were cloned by limiting dilution to ensure monoclonality (9). V gene fingerprinting analyses (11, 12) were performed to confirm that the V_H and V_L genes used by the heterohybridomas were derived from the CLL clones overexpanded in vivo. All seven hybridomas expressed V_H CDR3 lengths identical with those of the CLL cells (9). More recent studies indicated that six of the hybridomas had identical V_L CDR3 lengths. Since the fingerprinting results for hybridoma no. 039 were equivocal, the DNA sequence of this gene was determined from the circulating CLL cells.

Reverse Transcriptase (RT) PCR. The primers for the V_H and V_L families and C_H and C_L amplification were as reported in references 13 and 14, respectively. Total RNA (500 ng) was reverse transcribed by use of avian myeloblastosis virus RT (GIBCO BRL, Gaithersburg, MD) and the appropriate C_H or C_L primer, and then specific amplification of the cDNA was accomplished as described (15).

Cloning and Sequencing of PCR Products. PCR products were force cloned into the pUC19 vector after digestion with EcoRI and HindIII. Clones were isolated and sequenced in both directions by the Sanger dideoxy chain termination method (16), either manually or by an automated sequenator (model 373A; Applied Biosystems, Inc., Foster City, CA) using the DyeDeoxy™ terminator kit (Applied Biosystems Inc.).

Southern Blot Analyses. Genomic DNA was amplified by PCR with selected primers, electrophoresed through 1% agarose gel, transferred onto nylon membranes (Separation, Inc,. Westboro, MA), and UV cross-linked as described previously (15). Blots then were hybridized with ^{32}P-labeled probes in hybridization solution (QuikHyb; Stratagene Inc., La Jolla, CA) according to the manufacturer's instructions. When appropriate, stripping was accomplished by incubation at 72°C for 1 h in 2.5 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, and 0.5 mM sodium pyrophosphate. After washing, blots were exposed overnight on photographic film (XAR5; Eastman Kodak Co., Rochester, NY).

Results

V_H Gene Segment Analyses

Fig. 1 lists the Ig V_H gene sequences determined. Computer comparisons (17) of these with their presumed germline counterparts indicated the following features.

Biased V_H Gene Usage. Only two gene families were used, with five patients' cells using V_H4 family genes and two V_H3 family genes. Of the 12 or more members of this moderately sized family (18), only 2 were represented in this group: $V_H4.21$ (n = 3; patients 001, 033, and 055) and $V_H4.18$ (n = 2; 039 and 057). Because of the small number of V_H3 genes found, no conclusions about gene frequency use can be drawn.

Different Degrees of Similarity with Germline Genes. Two of the V_H4 CLL samples (patients 039 and 057) exhibit a high degree of homology (>98.5%) with the presumed germline ancestor (Fig. 1), whereas the other three (001, 033 and 055) display considerably less identity (95, 92, and 91%, respectively). Similarly, one of the V_H3 cases (040) is highly homologous with a germline gene (>98.5%), whereas much less similarity (94%) is observed in the other (030). The three V_H4 CLL B cells that show significant differences from their ancestral gene (<95% homology; 001, 033, and 055) all use the V_H 4.21 gene. Those patients using the V_H 4.18 gene exhibit >98.5% homology to the germline.

Differences in Distribution of Nucleotide Changes. For certain patients (001 and 033), the differences from the ancestral genes are located throughout the V_H segment (Fig. 1), whereas for others (030 and 055), differences are nonrandomly distributed. In these latter two patients, the differences cluster predominantly within complementarity determining region (CDR)2 for patient 030 and within CDR2 and FR3 for patient 055.

Differences in Deduced Amino Acid Sequences. When the nucleotide differences of the patients are analyzed for their effects on the amino acid sequence of the deduced Ig H chain proteins (Fig. 2 *A*), there are striking differences in replacement/silent (R/S) substitution ratios for certain patients. For example, for patient 030, the R/S of CDR 1,2 is 7.0 and that of framework regions (FR) 1,2,3 is 0.6. In contrast, for patient 033, the R/S in CDR 1,2 is 0.8 and in FR 1,2,3 is 1.3. Patient 055 illustrates an overall R/S for CDR 1,2 of 1.7 and for FR 1,2,3 of 1.3, although most of these latter substitutions occur in a defined region of FR3.

```
A   VH4

            Framework 1                                                                                                           CDR 1
         1   2   3   4   5   6   7   8   9   10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30   31  32  33  34  35
4.21     CAG GTG CAG CTA CAG CAG TGG GGC GCA GGA CTG TTG AAG CCT TCG GAG ACC CTG TCC CTC ACC TGC GCT GTC TAT GGT GGG TCC TTC AGT  GGT TAC TAC TGG AGC
CLL001   ... .G. ... ... ... ... ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..T ... ...  ... ... .T. ... ...
CLL033   ... ... ... ... ... ... ... ... A.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..T ... ... .A. ... ... ...  ... ..T ... ... .CT
CLL055   ... ... ... T.. ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... AT. ... ... ... ...  ... ..T C.. ... ..T

            Framework 2                                         CDR 2
         36  37  38  39  40  41  42  43  44  45  46  47  48  49   50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65
4.21     TGG ATC CGC CAG CCC CCA GGG AAG GGG CTG GAG TGG ATT GGG  GAA ATC AAT CAT AGT GGA AGC ACC AAC TAC AAC CCG TCC CTC AAG AGT
CLL001   ... ... ... .G. ... ..C ... ... ... ... ... ... ... ..T  ... ... ... ..C ... ..G .T. ... ... ... ... T.. ... ... ... ...
CLL033   ... G.. ... ... T.. ..C ..A ... ... ... ... ... ... ..A  C.. ... ... G.. ... ... ... ... ..T ..T ..T ... ... ... ... ...
CLL055   ... ... ... ... ... ..C ... ... ..T ... ... ... ... ...  ..G G.. .G. ... .C. ... .CT ... ... ... ... ... ... ... ... ...

            Framework 3
         66  67  68  69  70  71  72      73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97
4.21     CGA GTC ACC ATA TCA GTA GAC     ACG TCC AAG AAC CAG TTC TCC CTG AAG CTG AGC TCT GTG ACC GCC GCG GAC ACG GCT GTG TAT TAC TGT GCG AGA
CLL001   ... C.. ... T.. ... ... ...     ... ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..T ... ... ..G
CLL033   ..C ... ... ... ... ... ...     .A. ... ... ... ..A ... ... ... ..T ... ..G ... .C. ... ... ... ... ... ... ... ... ... ... ...
CLL055   ... ... ... ..G ... ... ... AGC .TT C.A .T. ... ... ... ... ... ... T.. .C. ... ... ... T.. ... ... ... ..T ... ... ... ... ...

            Framework 1                                                                                                           CDR 1
         1   2   3   4   5   6   7   8   9   10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30   31  32  33  34  35
4.18     CAG CTG CAG CTG CAG GAG TCG GGC CCA GGA CTG GTG AAG CCT TCG GAG ACC CTG TCC CTC ACC TGC ACT GTC TCT GGT GGC TCC ATC AGC  AGT AGT AGT TAC TAC
CLL039   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..T  ... ... ... ... ...
CLL057   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...  ... ... ... ... ...

                     Framework 2                                          CDR 2
         36  37   38  39  40  41  42  43  44  45  46  47  48  49  50  51   52  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67
4.18     TGG GGC  TGG ATC CGC CAG CCC CCA GGG AAG GGG CTG GAG TGG ATT GGG  AGT ATC TAT TAT AGT GGG AGC ACC TAC TAC AAC CCG TCC CTC AAG AGT
CLL039   ... ...  ... ... ... ... ... ... ... ... ... ... ... ... ... ...  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
CLL057   ... ...  ... ... ... ... ... ... ... ... ..T ... ... ... ... ...  ... ... ... ... ... ... ... ... ... ... ... A.. ... ... ... ...

            Framework 3
         68  69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99
4.18     CGA GTC ACC ATA TCC GTA GAC ACG TCC AAG AAC CAG TTC TCC CTG AAG CTG AGC TCT GTG ACC GCC GCA GAC ACG GCT GTG TAT TAC TGT GCG AGA
CLL039   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..C
CLL057   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

B   VH3

            Framework 1                                                                                                           CDR 1
         1   2   3   4   5   6   7   8   9   10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30   31  32  33  34  35
H11      GAG GTG CAG CTG GTG GAG TCC GGG GGA GGC TTA GTT CAG CCT GGG GGG TCC CTG AGA CTC TCC TGT GCA GCC TCT GGA TTC ACC TTC AGT  AGC TAC TGG ATG CAC
CLL030   ... ... ... ... ... ... ... ... ... .A. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...  .A. ... ... ... ...

            Framework 2                                         CDR 2
         36  37  38  39  40  41  42  43  44  45  46  47  48  49   50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66
H11      TGG GTC CGC CAA GCT CCA GGG AAG GGG CTG GTG TGG GTC TCA  CGT ATT AAT AGT GAT GGG AGT AGC ACA ACG TAC GCG GAC TCC GTG AAG GGC
CLL030   ... ... ... ... ... ... .A. ... ... .C. ... ... ... ..C  A.. ... .GA .C. ... ... G.C ... ... .AC ... ... ... G.. ... ... ...

            Framework 3
         67  68  69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98
H11      CGA TTC ACC ATC TCC AGA GAC AAC GCC AAG AAC ACG CTG TAT CTG CAA ATG AAC AGT CTG AGA GCC GAG GAC ACG GCT GTG TAT TAC TGT GCA AGA
CLL030   ... ... ... ... ... ... ... ..T ... ... ... ..A ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ...

            Framework 1                                                                                                           CDR 1
         1   2   3   4   5   6   7   8   9   10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30   31  32  33  34  35
1.9III   CAG GTG CAG CTG GTG GAG TCT GGG GGA GGC GTG GTC CAG CCT GGG AGG TCC CTG AGA CTC TCC TGT GCA GCC TCT GGA TTC ACC TTC AGT  AGC TAT GGC ATG CAC
CLL040   ... ... ... ... ... C.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ... ...  ... ... ... ... ...

            Framework 2                                         CDR 2
         36  37  38  39  40  41  42  43  44  45  46  47  48  49   50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66
1.9III   TGG GTC CGC CAG GCT CCA GGC AAG GGG CTG GAG TGG GTG GCA  GTT ATA TCA TAT GAT GGA AGT AAT AAA TAC TAT GCA GAC TCC GTG AAG GGC
CLL040   ... ... ... ... ... ... ... ... ... ... ... ... ... ...  ... ... .GG ... ... ... ... ... ... ... ... ... ... ... ... ... ...

            Framework 3
         67  68  69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98
1.9III   CGA TTC ACC ATC TCC AGA GAC AAT TCC AAG AAC ACG CTG TAT CTG CAA ATG AAC AGC CTG AGA GCT GAG GAC ACG GCT GTG TAT TAC TGT GCG AAA
CLL040   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ... ... ... ... ... .G.
```

Figure 1. Nucleotide sequences of V_H genes. The V_H germline genes most homologous to the individual CLL V_H genes are listed on top. These sequence data are available from EMBL/GenBank/DDBJ under accession numbers X84333-X84339.

Absence of the Patient-specific Sequences in Germline DNA

Next we investigated whether the differences in the V_H gene segments observed in patients 030 and 055 were somatically generated. These patients were chosen since their V_H sequences were significantly different from the presumed germline counterparts, and these differences clustered within areas of the gene that would likely directly influence the shape of the antigen-binding groove.

PCR Analyses. The first approach involved the construction of two sets of PCR primer pairs, one that would allow amplification of a region encompassing the entire CDR2 and portions of FR2 and FR3 of all genes of the two families involved (V_H3 and V_H4), and another set for amplification of only the unique sequences from the two individual patients (CDR2 for patient 030 and CDR2 plus part of FR3 for patient 055). The former are designated *G* for general primers, and the latter *S* for specific primers. These primer pairs were used to amplify either unrearranged neutrophil DNA or rearranged lymphocyte DNA. Fig. 3 indicates the results obtained for patient 055.

Fig. 3, Lanes *1* and *2* indicate the positions in the gel of

A H CHAIN

```
                   Framework 1           CDR 1 Framework 2   CDR 2                    Framework 3                   CDR 3                    Framework 4
4.21  QVQLQQWGAGLLKPSETLSLTCAVYGGSFS G  YYWS WIRQPPGKGLEWIG EINHSGSTNYNPSLKS  RVTISVD TSKNQFSLKLSSVTAADTAVYYCAR
CLL001.G............................ .  .F.. ...R.......... ......I....S....  .L.L... .........................WYYFDTSGYYPRNFYYMDV      WGKGTPVTVSS
CLL033.......T.................E.... .  ...T .V..S......... Q..D............  ....... K.......N.R.A............ SRFYCSGETCHSSQFYYYHGLDA WGQGTTVTVSS
CLL055.......................I...... .  .H.. .............. .VS.T.T.........  ...M...SIPM.......T....S.........APLGGGAGLYNWFDP          WGLGTLVTVSS

4.18  QLQLQESGPGLVKPSETLSLTCTVSGGSIS SSSYYWG WIRQPPGKGLEWIG SIYYSGSTYYNPSLKS  RVTISVD TSKNQFSLKLSSVTAADTAVYYCAR
CLL039.............................. ....... .............. ................  ....... ........................SSRGYSSSWWSS NWFDP       WGQGTLVTVSS
CLL057.............................. ....... .............. ...........T....  ....... .........................HLGYSSSWYGAANWFDP        WGQGTLVTVSS

H11   EVQLVESGGGLVQPGGSLRLSCAASGFTFS SYWMH   WVRQAPGKGLVWVS RINSDGSSTTYADSVKG RFTISRD NAKNTLYLQMNSLRAEDTAVYYCAR
CLL030.........D.................... N....   ......E..P.... S.RT..G..N...A... ....... .........................AHSPHGSHYPS              WGQGTLVTVSS

1.9IIIQVQLVESGGGVVQPGRSLRLSCAASGFTFS SYGMH   WVRQAPGKGLEWVA VISYDGSNKYYADSVKG RFTISRD NSKNTLYLQMNSLRAEDTAVYYCAK
CLL040.....Q........................ .....   .............. ..W.............. ....... ........................RDRGIGGWQNYMDV            WGKGTTVTVSS
```

B L CHAIN

```
              Framework 1             CDR 1            Framework 2       CDR 2        Framework 3                    CDR 3        Framework 4
DPL8    QSVLTQPPSVSGAPGQRVTISC   TGSSSNIGAGYDVH   WYQQLPGTAPKLLIY   GNSNRPS  GVPDRFSGSKSGTSASLAITGLQAEDEADYYC  QSYDSSLSG
CLL001  .............A..G.....   ...N..V.......   .......K....V.H   .......  ...G........A.........S........  .T...R.NICV  FGGGTRLTVVG

L12a    DIQMTQSPSTLSASVGDRVTITC  RASQSISSWLA      WYQQKPGKAPKLLIY   KASSLES  GVPSRFSGSGSGTEFTLTISSLQPDDFATYYC  QQYNSYS
CLL030  .......................  .....VT.RM.      ...............   RG...Q.  ...L............................  ...LN.PVT    FGGGTKVEIXR

A27     EIVLTQSPGTLSLSPGERATLSC  RASQSVSSSYLA     WYQQKPGQAPRLLIY   GASSRAT  GIPDRFSGSGSGTDFTLTISRLEPEDFAVYYC  QQYGSSP
CLL033  ...........KV..........  ........T...     ............V..   .T.....  ................................  .....L.ST    FGRGTKVEIK

L9      AIRMTQSPSSFSASTGDRVTITC  RASQGISSYLA      WYQQKPGKAPKLLIY   AASTLQS  GVPSRFSGSGSGTDFTLTISCLQSEDFATYYC  QQYYSYP
CLL040  .......................  ...........      ...............   .......  ................................  .......QT    FGQGTKVEIK

L19     DIQMTQSPSSVSASVGDRVTITC  RASQGISSWLA      WYQQKPGKAPKLLIY   AASSLQS  GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC  QQANSFP
CLL055  .....................S.  ....H..T...      ....R.........S   G.....G  .............H.....TG.........F.  .......IT    FGGGTKVEIFP

O2      DIQMTQSPSSLSASVGDRVTITC  RASQSISSYLN      WYQQKPGKAPKLLIY   AASSLQS  GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC  QQSYSTP
CLL039  .......................  ...........      ...............   .......  ................................  .......RS    FGQGTKLEIKR
CLL057  .......................  ...........      ...............   .......  ................................  .......RT    FGRGTKVEIKR
```

Figure 2. Deduced amino acid sequences of the H and L chains. The amino acid sequence of the germline gene is given through FR 3, with the CLL sequences listed below. Sequences of each CLL CDR3 and FR4 are presented. (*A*) H chain. (*B*) L chain.

the V_H 4.21 segment amplified from RNA by RT/PCR with the G and S primer pairs, respectively. Amplification of both germline (lane *3*) and rearranged (lane *5*) DNA with the G primer set resulted in appropriately sized products. Amplification of the same DNA with the S primer pair (lanes *4* and *6*) failed to reveal readily detectable bands. Since the inability to consistently amplify a product may be due to gene copy number, aliquots (1 μl) of the products of the G primer amplifications from both unrearranged and rearranged DNA, which should have increased specifically the frequency of all V_H4 CDR2 regions, were reamplified with the S primers. After reamplification, only the rearranged DNA yielded an appropriate product (lane *8*). Products corresponding in size to the specific mutated sequence never were detected with germline DNA (lane 7). Faint bands corresponding to the G products can be seen for each sample (lanes 7 and *8*) because of the transfer of G primers from the original reactions of lanes *3* and *5*. Most importantly, DNA sequence analyses of the isolated G and S bands from lanes 7 and *8*, respectively, revealed identity with those presented in Fig. 1 (data not shown).

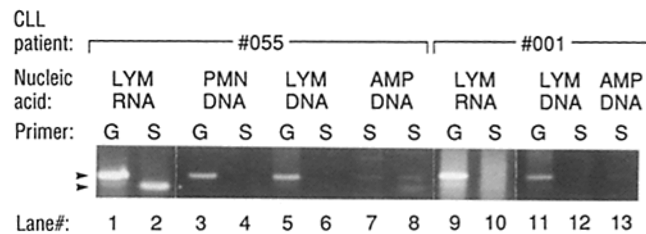

Figure 3. Absence of CLL 055–specific sequences in germline DNA as assessed by PCR. Gel indicating PCR fragments amplified from various nucleic acids by use of the general primer pairs (G: 55.G.Fwd.: 5'AGGGTCTGGAGTGGATTGGG and 55.G.Rev. 5'ACTTCAGGGAGAACTGGTTC) and specific (S: 55.S.Fwd. 5'GAGGTCAGTCATACTGGAAC and 55.S.Rev. 5'GAACTGGTTCATTGGAATGC). Note the absence of a CLL 055–specific fragment in lane 7 containing the products of the reamplified neutrophil DNA from lane *3*. In contrast, an appropriately sized product is detected in lane *8* after reamplification of the CLL DNA from lane *5*.

As a further specificity control, these primers were used to amplify the G and S sequences from another patient whose CLL B cells used the same Ig V_H gene (4.21; patient 001). An appropriately sized product was obtained from rearranged DNA with the G primers (lane *11*), but not with the S primers (lane *12*). In addition, products could not be found when the G product was reamplified with the S primers (lane *13*), indicating that the S sequence was uniquely expressed only in CLL 055 B cells and not in a patient using the same ancestral V_H4 gene. Identical results were obtained by performing similar studies with patient 030 (data not shown).

Southern Blotting Analyses. A second approach involved the probing of unrearranged DNA with radiolabeled oligonucleotides specific for the mutation-specific or conserved se-

quences of each patient. In these studies, germline and rearranged DNA were amplified with the same sets of general G primers described above and then exposed to probes complementary to either the unique CDR2 sequence or the conserved FR sequence as a positive control. Fig. 4 indicates the results of these experiments for patient 030. Hybridization with the mutation-specific probes occurred only with the rearranged DNA, whereas hybridization of the conserved probe was seen with both the unrearranged and rearranged DNA. These differences in binding of the mutation-specific probes occurred despite the use of longer incubation times (16 vs. 2 h) and less stringent hybridization buffers and washing regimens (2 vs. 0.1 × SSC; room temperature vs. 40°C) than those used for the conserved probes.

D and J_H Gene Segment Analyses

The sequences of the D and J_H segments expressed by the $CD5^+$ CLL B cells are listed in Fig. 5, *A* and *B*.

Nonrandom J_H Gene Usage. J_H5 and J_H6 gene segments were used by six of the seven patients. Surprisingly, there were no examples of J_H4 gene use in this cohort, even though this is the most commonly used J_H segment in healthy adults (19). D gene use appeared random, although homologous germline D segments could not be accurately defined in two patients (030 and 055).

Somatic Diversification of D and J_H Segments. Among those CLL B cells with clearly definable D segments, varying degrees of identity with the germline gene segments existed (Fig. 5 *A*), with patients 001 and 057 being the most homologous (92.8 and 90.5%) and patients 033, 039, and 040 being the least (67.7, 58.8, and 47.7%, respectively). Ancestral germline D segments could not be assigned for patients 030 and 055. Indeed, in these cases, it appeared that the majority (030) or the entire (055) D segment was deleted at the time of recombination and replaced by N addition. It is noteworthy that the V_H genes of these patients were two of the most divergent from their germline counterparts (Fig. 1).

The J_H segments also show significant alterations from the germline, with examples of both trimming at the D–J joints and internal deletion (Fig. 5 *B*). Furthermore, significant numbers of N additions were identifiable for all patients except 040.

Finally, presumptive nucleotide substitutions were found in the J_H and D segments of all patients in whom likely germline progenitor D segments could be identified. Patients 039 and 040 are noteworthy since in both cases the level of V_H gene diversification was negligible, whereas the levels of DJ_H mutation were considerable.

Deduced Amino Acid Differences in CDR3. The nucleotide changes in the D and J_H segments resulted in CDR3 lengths that were quite disparate, ranging from 11 to 23 amino acids (Fig. 2 *A*). A dramatic similarity in amino acid sequence was found in the CDR3 of patients 039 and 057. This sequence similarity includes an identical stretch of 6 amino acids that is separated by four mismatches from another identical stretch of 5 amino acids.

Reidentification of CLL-unique Sequences in Freshly Isolated $CD5^+$ CLL B Cells

Since this study spanned 3 yr, attempts were made to determine if new mutations had occurred in vivo over this time frame. Patients 001, 030, 033, 039, and 057 were reanalyzed. Patients 055 and 040 could not be included in this evaluation since the former was treated with fludarabine that essentially eliminated circulating CLL B cells, and the latter had died. The unique V gene arrays that had been created by mutation, N addition, and D/J segment substitution and deletion served as unequivocal signatures of the CLL clones.

RNA was obtained from freshly isolated B cells, reverse transcribed, amplified with appropriate patient-specific FR2 upstream primer and the Cγ downstream primer, and sequenced. In three patients (001, 033, and 039), the newly defined sequences were identical to those listed in Fig. 1 (data not shown). A single-point alteration that was not productive of an amino acid change was found in patient 030. The most significant differences were found in patient 057, in whom three point mutations in CDR 2 and 3 were detected among two clones; one of these (G → A) resulted in an amino acid change (Ser → Asn) at codon 52 of CDR2.

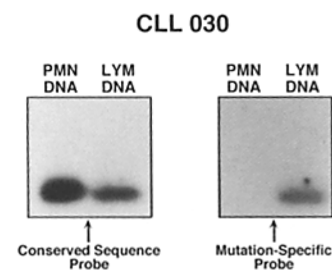

Figure 4. Absence of CLL 030–specific sequences in germline DNA as assessed by Southern blotting. Genomic DNA from either B cells or neutrophils was amplified by PCR by use of the general primer pairs for patient 030 (G: 30.G.Fwd. CCGCCAAGCTCCAGAGAAGG and 30. G.Rev. 5'GGCATTGTCTCTGGAGATGG) and transferred onto nylon membranes. Blots were hybridized first with a mutation-specific ^{32}P-labeled probe (5'CAGTATTAGAACTGATGGGG), and after stripping by incubation at 72°C for 1 h in 2.5 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, and 0.5 mM sodium pyrophosphate, were hybridized with the conserved sequence probe (5'GTGAAGGGCCGATTCA). For the mutation-specific probe, hybridization was carried out at 42°C overnight; for the conserved sequence probe, hybridization was performed at the same temperature for 2 h. For the conserved sequence probe, blots were washed twice at room temperature for 15 min with 2 × SSC, 0.1% SDS, and then once at 40°C for 30 min with 0.1 × SSC, 0.1% SDS. For the mutation-specific probe, washes were performed under less stringent conditions with 2 × SSC, 0.1% SDS, at room temperature for 15 min.

V_L Gene Segment Analyses

Fig. 6 lists the Ig V_L gene sequences determined and the computer comparisons with the presumed germline counterparts.

Nonrandom V_L Gene Family Use. Of the six CLL clones expressing κ chains, five used a $V_\kappa I$ family gene (patients 030, 040, 055, and 057). However, unlike the Ig V_H4 gene use, which was restricted to only two individual family members, four different $V_\kappa I$ family genes were represented (Fig. 6 and Table 1).

Different Degrees of Similarity with Germline V_L Genes That Parallel Those of the Corresponding V_H Gene. The $V_\kappa I$ genes

A D GENE SEGMENTS AND N REGION ADDITIONS

Patient	FR3	N	D Segment	N
			DLR2	
			AGGATATTGTAGTGGTGGTAGCTGC TACTCC	
CLL033	GCG	TCTCGCTTT	.............AA.CT... C....A	TCACAGTTTT
CLL040	GCGAGA		G..A.GG.AAT....G....GCAAA...A.	
			D21/9	
			GTATTACTATGATAGTAGTGGTTATTAC	
CLL001	GCGAGG	TG	........T.....C.............	CCCCG
			DM1	
			GGTATAACTGGAACTAC	
CLL039	GCGAGC	TCCAGAG	......G.A.C.GC.GG	TGGTCATCT
			DM1	
			GGGTATAGCAGCAGCTGGTAC	
CLL057	GCGAGA	CATCTG	..A.................T	GGGGCAGCC
			UNASSIGNABLE	
CLL030	GCAAGA		GCTCACTCGCCTCACGGCAGCCACTATCC	
CLL055	GCGAGA		GCCCCTCTGGGGGGTGGGGCGGGGCTTT	

B J_H GENE SEGMENTS

Patient	CDR 3	Framework 4
J_H2	TACTGGTACTTCGATCTC	TGGGGCCGTGGCACCCTGGTCACTGTCTCCTCA
CLL039	A...... ....C.C.	.......AG..A...........C.........
J_H5	ACAACTGGTTCGACTCC	TGGGGCCAAGGAACCCTGGTCACCGTCTCCTCA
CLL057	............C..	........G........................
CLL055	..............C..	.......TG................T......
CLL030	...T	........G........................
J_H6	TACTACTACTACTACGGTATGGACGTC	TGGGGGCAAGGGACCACGGTCACCGTCTCCTCA
CLL001	.A...T....TAC.........	.....CA.G......C.................
CLL033	........C.....T......C.	.....C..G........................
CLL040	....	.....CA..........................

Figure 5. Nucleotide sequences of the CDR3 and FR4 regions of the H chains. (*A*) Sequences of the CLL D gene segments compared with their most likely germline counterpart. Those listed as "unassignable" were not sufficiently similar to 24 reported sequences to make a comparative ancestral assignment. (*B*) Comparison of germline and CLL J_H gene segments.

used by patients 040, 039, and 057 displayed complete identity with the germline genes L9 and O2, respectively, and therefore are not listed on Fig. 6. In contrast, the V_L genes used by patients 001, 030, 033, and 055 exhibited ≥5% divergence from the most homologous germline counterparts. These levels of difference parallel those seen among the Ig V_H genes of each patient.

Differences in Distribution of Nucleotide and Amino Acid Changes Also Parallel Those of the Corresponding V_H Gene. For patients 001 and 033, nucleotide changes were distributed throughout the V_L segment, whereas for patients 030 and 055 these clustered either within CDR 1, 2, and 3 for patient 030 or in CDR 1, 2, and 3 and FR 2 for patient 055. The deduced amino acid V_L sequence of patient 030 exhibits

Framework 1 (1–23) / CDR 1 (24–34)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	27A	27B	27C	28	29	30	31	31A	32
DPL8	CAG	TCT	GTG	CTG	ACG	CAG	CCG	CCC	TCA		GTG	TCT	GGG	GCC	CCA	GGG	CAG	AGG	GTC	ACC	ATC	TCC	TGC	ACT	GGG	AGC	AGC	TCC	AAC	ATC	GGG	GCA	GGT	TAT		GAT
CLL001	...	...	...	...	...	...	...	...	...		...	...	...	...	G..	...	...	G..	...	...	...	...	...	...	...	...	.A.	...	...	G..	...	...	...	...		...
L12a	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCT	TCC	ACC	CTG	TCT	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	CGG	GCC	AGT	CAG				AGT	ATT	AGT	AGC		TGG
CLL030	...	...	...	...	...	...	...	..C	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...				...	G..	.C.	...		C..
A27	GAA	ATT	GTG	TTG	ACG	CAG	TCT	CCA	GGC	ACC	CTG	TCT	TTG	TCT	CCA	GGG	GAA	AGA	GCC	ACC	CTC	TCC	TGC	AGG	GCC	AGT	CAG				AGT	GTT	AGC	AGC	AGC	TAC
CLL033	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	.A.	.T.	...	...	...	...	...	...	...	...				...	...	..T	..T	.C.	...
L19	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCT	TCT	GTG	TCT	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	CGG	GCG	AGT	CAG				GGT	ATT	AGC	AGC		TGG
CLL055	...	...	...	...	...	...	...	...	...	...	...	...	..T	...	...	...	...	...	...	...	...	.G.	...	...	...	...	..A				CA.	...	...	.C.		...

Framework 2 (35–49) / CDR 2 (50–56) / Framework 3 (57–88)

	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68
DPL8	GTA	CAC	TGG	TAC	CAG	CAG	CTT	CCA	GGA	ACA	GCC	CCC	AAA	CTC	CTC	ATC	TAT	GGT	AAC	AGC	AAT	CGG	CCC	TCA	GGG	GTC	CCT	GAC	CGA	TTC	TCT	GGC	TCC	AAG	TCT	GGC
CLL001	..G	...	...	...	...	...	...	...	...	.A.	...	...	...	...	G..	...	C..	...	..T	...	...	...	...	...	...	...	...	.G.	...	...	...	...	...	...	...	...
L12a	TTG	GCC	TGG	TAT	CAG	CAG	AAA	CCA	GGG	AAA	GCC	CCT	AAG	CTC	CTG	ATC	TAT	AAG	GCG	TCT	AGT	TTA	GAA	AGT	GGG	GTC	CCA	TCA	AGG	TTC	AGC	GGC	AGT	GGA	TCT	GGG
CLL030	A..	...	...	...	...	...	...	...	...	...	...	...	..A	...	...	...	...	.G.	.GA	..C	...	...	C..	...	...	...	...	CT.	...	...	...	...	...	...	...	...
A27	TTA	GCC	TGG	TAC	CAG	CAG	AAA	CCT	GGC	CAG	GCT	CCC	AGG	CTC	CTC	ATC	TAT	GGT	GCA	TCC	AGC	AGG	GCC	ACT	GGC	ATC	CCA	GAC	AGG	TTC	AGT	GGC	AGT	GGG	TCT	GGG
CLL033	...	...	...	...	..A	...	...	...	...	...	...	...	...	...	G..	..T	...	...	A..	...	...	...	...	..C	...	...	...	...	...	...	...	...	...	...	...	...
L19	TTA	GCC	TGG	TAT	CAG	CAG	AAA	CCA	GGG	AAA	GCC	CCT	AAG	CTC	CTG	ATC	TAT	GCT	GCA	TCC	AGT	TTG	CAA	AGT	GGG	GTC	CCA	TCA	AGG	TTC	AGC	GGC	AGT	GGA	TCT	GGG
CLL055	...	...	...	...	...	...	.G.	...	..A	...	...	...	..A	...	..T	...	.C.	.G.	...	...	...	...	...	G..	...	...	...	...	...	...	...	...	...	...	...	...

Framework 3 (69–88) / CDR 3 (89–95B)

	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	95A	95B
DPL8	ACC	TCA	GCC	TCC	CTG	GCC	ATC	ACT	GGG	CTC	CAG	GCT	GAG	GAT	GAG	GCT	GAT	TAT	TAC	TGC	CAG	TCC	TAT	GAC	AGC	AGC	CTG	AGT	GGT
CLL001	G..	...	...	...	...	...	...	...	..A	...	...	T..	...	...	...	...	...	...	...	...	...	A..	...	...	...	..G	...	.A.	AT.
L12a	ACA	GAA	TTC	ACT	CTC	ACC	ATC	AGC	AGC	CTG	CAG	CCT	GAT	GAT	TTT	GCA	ACT	TAT	TAC	TGC	CAA	CAG	TAT	AAT	AGT	TAT	TCT		
CLL030	...	...	...	...	...	...	...	...	...	...	...	...	...	..C	...	...	...	...	...	...	...	...	..C	CT.	.A.	..C	C.G		
A27	ACA	GAC	TTC	ACT	CTC	ACC	ATC	AGC	AGA	CTG	GAG	CCT	GAA	GAT	TTT	GCA	GTG	TAT	TAC	TGT	CAG	CAG	TAT	GGT	AGC	TCA	CCT		
CLL033	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	..A	...	...	...	..T	.T.	...		
L19	ACA	GAT	TTC	ACT	CTC	ACT	ATC	AGC	AGC	CTG	CAG	CCT	GAA	GAT	TTT	GCA	ACT	TAC	TAT	TGT	CAA	CAG	GCT	AAC	AGT	TTC	CCT		
CLL055	...	C..	...	...	...	...	...	.C.	G..	...	...	...	...	...	...	...	...	...	.T.	...	...	..A	...	...	...	...	..G		

Figure 6. Nucleotide sequences of V_L genes. Since the V_L gene sequence for patient 040 is identical to the L9 germline gene and those of patients 039 and 057 are identical to the O2 germline gene, they are not listed. These sequence data are available from EMBL/GenBank/DDBJ under accession numbers X84340-X84346.

Table 1. *Serologic and Genetic Characteristics of IgG⁺CD5⁺ CLL B Cells**

						H chain				
								Types of diversification		
Patient	C_H	CRI‡	V_H family	Most homologous germline V_H gene	V_H gene difference	Likely germline D segment	J_H	N addition/ trimming§	Somatic‖ mutation V_H/DJ_H	Possible antigenic selection
					%					
CLL 001	γ1	–	4	4.21	5	D21/9	6	+/+	++/++	–
CLL 033	γ1	F4	4	4.21	8	DLR2	6	+++/+	+++/+++	–
CLL 055	γ1	F4	4	4.21	9	?	5	+++/–	+++/+	+/–
CLL 039	γ3	–	4	4.18	1	DM1	2	++/–	–/+++	+
CLL 057	γ3	–	4	4.18	1	DN1	5	++/+	–/+	+
CLL 030	γ1	F4	3	H11	6	?	5	+++/+++	++/+	+
CLL 040	γ1	–	3	1.9III	2	DLR2	6	–/+++	–/+++	+

						L chain				
								Types of diversification		
Patient	C_L	CRI‡	V_L family	Most homologous germline V_L gene	V_L gene difference		J_L	N addition/ trimming§	Somatic‖ mutation V_L/J_L	Possible antigenic selection
					%					
CLL 001	λ	–	λI-c	DPL8	6		Jλ2/3	–/+	++/+	–
CLL 033	κ	3I	κIIIb	A27	5		Jκ1	–/+	++/+	–
CLL 055	κ	3I	κI	L19	7		Jκ4	+/+	+++/+	+/–
CLL 039	κ	–	κI	O2	0		Jκ2	+/++	–/+	+
CLL 057	κ	–	κI	O2	0		Jκ1	+/+	–/+	+
CLL 030	κ	3I	κI	L12a	7		Jκ4	–/+	+++/+	+
CLL 040	κ	3I	κI	L9	0		Jκ1	–/+	–/+	–

* Adapted, in part, from references 9 and 22.
‡ Cross-reactive idiotypic (CRI) determinants detected by solid-phase RIA as described (33, 34).
§ Notations preceding the slash (/) relate to the relative numbers of N additions; notations after the slash relate to degrees of trimming at joints.
‖ Notations preceding the slash relate to the relative number of mutations in V_H or V_L; those after the slash relate to DJ_H and J_L segments.

R/S substitution ratios >3 for the CDR and <1 for the FR (Fig. 2 *B*). These patterns closely resemble those seen for the corresponding H chain gene in almost every case.

Random Use of J_L Gene Segments. Unlike the nonrandom use of J_H, J_L families are used in an apparently stochastic fashion (Fig. 7). In every case, there is evidence for trimming of the J_L segments as well as examples of a limited number of N additions in patients 039, 040, and 057 (Fig. 7). The combination of trimming and N addition has resulted in V_L–J_L junctional diversity in each case.

Discussion

The preceding data demonstrate that certain CD5⁺ CLL B cells display skewed use, somatic diversification, and evidence consistent with antigenic selection of their Ig H and L chain V region genes. Table 1 collates the genetic data from this study and the serologic data from a companion study (9) to provide a comprehensive view of the characteristics of these cells.

Gene Segment Use. Evidence for nonrandom use of both H and L chain V region genes was manifested as an overrepresentation of V_H4 and $J_\kappa I$ family genes and underrepresentation of the J_H4 gene segment. V_H4 gene use was restricted further to only two (4.21 and 4.18) of the 12–16 members of the family (18). A restriction to specific gene members was not seen for $J_\kappa I$, although the O2 gene was expressed in two patients.

The frequency of V_H4 gene use among our patients is

```
                            CDR3                                  Framework 4
Patient     |-----VL----|    N      |---------------------JL---------------------|

                                     TG TG  GTA     TTC GGC GGA GGG ACC AAG CTG ACC GTC CTA G  [Jλ2/3]
CLL 001     ... ... AAT ATT  -         ..C ..G      ... ... ..G ... ... .G. ... ... ... G.. .

                                      G TGG ACG     TTC GGC CAA GGG ACC AAG GTG GAA ATC AAA C  [Jκ2]
CLL 033     ... ... TTA CCT  -         .C.  ...     ... ... .G. ... ... ... ... ... ... ... .
CLL 040     ... ... TAC CCT  CA          .  ...     ... ... ... ... ... ... ... ... ... ... .
CLL 057     ... ... ACC CCT  C          .A  ...     ... ... ... ... ... ... ... ... ... ... .

                                     TG TAC ACT     TTT GGC CAG GGG ACC AAG CTG GAG ATC AAA C  [Jκ2]
CLL 039     ... ... ACC CCT  CGG            .G.     ... ... ... ... ... ... ... ... ... ... .

                                      G CTC ACT     TTC GGC GGA GGG ACC AAG GTG GAG ATC AAA C  [Jκ4]
CLL 030     ... ... TAC CCG  -        G.. ...       ... ... ... ... ... ... ... ... ... ... .
CLL 055     ... ... TTC CCG  -        A.. ..C       ... ... ... ... ... ... ... ... ... ... .
```

Figure 7. Comparison of the nucleotide sequences of the J_L gene segments.

significantly different from the potential repertoire available in the germline (exact goodness of fit test; $p < 0.005$; 20) but not from a composite of the available CLL V gene sequences (21). Nevertheless, the overexpression of the specific V_H4 genes 4.21 and 4.18 is striking considering that two prior studies failed to document such a restricted use of individual genes within a V_H family in CLL B cells (23, 24). Furthermore, the absence of J_H4 use by our patients' B cells is significantly different not only from healthy adult blood B cells ($p < 0.003$; 19) but also from most CLL patients ($p < 0.002$; 21). In addition, the overrepresentation of $V_\kappa I$ family genes in our patients differs from that of the $V_\kappa IIIb$ gene humkv325/A27 reported in other cohorts with CLL (25), although one of our patient's (033) CLL cells did use this gene. Finally, the biased use of Ig genes by our CLL patients extends to C_H and involves the restricted use of IgG1 and IgG3 and the absence of IgG2 (9, and Table 1). The lack of IgG2 expression is striking since in healthy adults, B cells producing this subclass are of equal frequency to those producing IgG1 (26).

Thus our data support the principle that $CD5^+$ CLL B cells demonstrate biased use of Ig H and L chain gene segments. Those biases observed in our patients that differ from those reported previously may indicate different selective pressures in these IgG-producing CLL cells compared with most CLL patients that develop a clonal expansion of IgM-bearing B cells (92–95%; 27). These findings favor the concept that the V gene biases in $CD5^+$ B cells can result from antigen selection (28, 29) rather than inherent processes programmed before antigen receptor expression (30).

Somatic Diversification. Some level of somatic diversification was found in every case. These diversification events occurred both at the time of gene segment recombination as well as after the attainment of a competent B cell antigen–receptor complex. For example, at the times of D–J_H and V_H–DJ_H joining and V_L–J_L joining, varying levels of N addition and trimming occurred in the H and L chains of most patients. It is noteworthy that the combined effects of trimming and N addition in V_L, however, conserved CDR3 lengths of nine amino acids. It is now known that ~20% of normal human B cells exhibit additions of nucleotides at the V_L–J_L junction that can alter CDR3 length (31), although longer CDR3 lengths appear to be more frequent in patients with inflammatory diseases like rheumatoid arthritis (32).

Multiple somatic changes not related to the recombination process also occurred in regions of these genes. Four patients showed significant variations of their H and L chain V genes from the reported germline counterparts. The level and location of these differences in the H and L chains were quite similar within an individual patient's leukemic clone. The possibility that these differences reflected the presence of heretofore unreported gene segments was ruled out for the two patients studied in detail (030 and 055). Interestingly, a cross-reactive idiotype found almost exclusively on mutated IgG autoantibodies and not within the germline (F4; 33) was expressed by these two IgG molecules as well as that produced by patient 033 (9 and Table 1), providing serologic corroboration of the molecular data obtained.

However, mutations were not seen in the V_H genes of all patients studied. Specifically the two patients using the V_H 4.18 gene demonstrated very minor differences from the germline counterpart. In contrast, the three patients that used the V_H 4.21 gene demonstrated significant somatic diversification. This observation confirms those recently reported (21, 35) on the discordant occurrence of mutations in these two gene segments in CLL.

The diversification detected in the IgG-producing B cells of these patients is not restricted to the V_H or V_L segments. Essentially all of the patients demonstrated differences from the germline D and J_H segments, and every patient showed diversity at the V_L–J_L junctions. These latter changes are striking since they result in the acquisition of a positively charged arginine residue at the joints in two patients (039 and 057) and the creation of a threonine residue one position downstream from the joints in five patients (030, 033, 040, 055, and 057).

Most previous reports suggest that $CD5^+$ CLL B cells undergo little somatic diversification (2), although certain studies have contained occasional patients whose V genes diverge from the presumed germline counterparts (for review

see reference 21). The most striking study of V_H gene mutation in CLL is that of Cai et al. (6), who documented extensive and selected mutations in IgM-producing CLL B cells that used the V_H251 gene. This group of patients was somewhat atypical, however, since ~30% of the patients studied used this gene segment, a frequency not in line with most other studies. Since neither the frequency of V_H251 use nor the presence of extensive mutations could be corroborated by Rassenti and Kipps (7), these investigators suggested that the patients studied by Cai et al. (6) might not be representative of the disease or that subsets of patients' CLL B cells might differ in their ability to develop somatic mutations. In support of the latter possibility are reports of V_H (36) and V_L (37) gene mutations in CLL patients whose B cells lacked CD5 expression. Although our patients display typical features of CLL and their expanded leukemic clones express CD5, they do comprise a subset defined by the production of IgG antibodies (9). Thus, our data are compatible with both the studies of Cai et al. and Rassenti and Kipps, providing clear evidence for somatic diversification of V region genes in a subset of leukemic $CD5^+$ B cells that have undergone an isotype class switch. Furthermore, our studies indicate that mutations can occur in the V genes of both the H and L chains and that the numbers and locations of these changes are closely paralleled in individual patients. However, it should be noted that the extent of mutation documented in these leukemic $CD5^+$ B cells is not as extensive as that observed in $CD5^+$ B cells from sites of chronic inflammation (38).

Antigen Selection. The diversification detected in this study is distributed differently among the V region genes of the various patients. In some cases (e.g., patients 001 and 033), nucleotide differences are randomly distributed, whereas in others (e.g., patient 030) these are clustered in the CDR. In contrast, patients 039, 040, and 057 demonstrate virtually no changes in their V_H or V_L segments, but exhibit extensive changes in D and J_H and unique changes at the V_L–J_L junctions. Although the H chain disparities could have arisen by V_H gene replacement, the lack of mutation seen in the corresponding V_L segments of these same patients suggests that this mechanism is a less likely explanation.

These somatic changes result in significant and differing numbers of replacement changes at the amino acid level. It is accepted that the presence of an R/S ratio ≥ 3 in the CDR with lower ratios in the FR suggests antigen selection of the somatically generated substitution (39). These criteria are met for the V_H of patient 030 (R/S >7 in CDR and 0.6 in FR). In addition, the clustering of replacement substitutions in CDR2 and between amino acids 73 and 76 of FR3 of patient 055 suggest antigen selection, since this portion of FR3 of the V_H is known to affect antibody binding to DNA (40, 41) and the conformation of the antigen-binding site (42). Similarly, the V_L of patient 030 has an R/S substitution ratio clearly compatible with antigen selection (CDR, 3.3; FR, 0.3), and the V_L of patient 055 has an R/S ratio and pattern similar to that in V_H, again with involvement of the amino acids in the CDR and FR3, at or around position 76 of the L chain.

However, the most striking indications of selection are the amino acid changes that have occurred in the CDR3 of the H and L chains of patients 039 and 057. Indeed, despite the use of different germline D and J_H genes, the addition of different numbers and types of nucleotides at the joints, and the difference in levels of trimming or deletion, identity at 11 of 17 amino acid positions has been achieved by somatic processes (Fig. 2 *A*). In addition, a basic arginine residue has been created at the V_L–J_L joints in each case (Fig. 2 *B*). This also is striking since these two B cells used different J_κ genes and exhibited different levels of trimming and N addition at these joints. Considering that both patients' CLL cells use the same, virtually nonmutated V_H (4.18) and V_L (O2) genes, the amino acid structure of their two antigen receptors is >95% similar. Thus, there appear to be selective pressures, presumably antigenic, to mutate the CDR3 of V_H and the V–J junctions of V_L to yield remarkably similar amino acid sequences and maintain the identity of the rest of the combining sites in these two patients. Finally, since both patients' CLL cells use $\gamma 3$ H chains, one must consider strongly that a similar type of antigenic drive and cellular help has been involved in their diversification and selection processes.

Implications for CLL and Its Evolution. These data demonstrate conclusively that somatic mutations of Ig V region genes can be found in certain CLL clones. However, the data do not indicate when these mutations occurred in relation to the leukemogenic event(s). If they occurred before, then our data would be consistent with the notion that transformation renders these cells/genes less "mutable." In addition, they would indicate that the transformation event need not occur in a relatively immature B cell whose V genes still have a germline-like sequence, as has been suggested recently (21). In contrast, if these mutations are ongoing, this would suggest that these cells behave more like their normal counterparts. Our 3-yr follow-up study of the accumulation of only limited numbers of mutations between FR2 → J_H in five patients is relevant here. However, this issue needs more extensive study in more patients.

Our data also are consistent with the idea that antigenic stimulation may be involved in the clonal amplification in this disease. Studies of other lymphoid malignancies are compatible with this view (43–45). Considering the known propensity of $CD5^+$ CLL cells to be autoreactive (3–5), autoantigenic drive through the B cell antigen receptor is an attractive possibility. However, surprisingly, the IgG antibodies secreted by these $CD5^+$ CLL cells did not react significantly with a panel of six classical autoantigens (9). This observation appears to relate at least in part to antibody valency (9). However, in addition, the somatic mutations documented here may have sufficiently altered their binding sites away from reactivity with an original (auto)antigen in favor of another or of an exoantigen (46, 47).

Finally, these studies suggest that our understanding of the triggering capabilities and requirements for CLL cells may be too simplistic. Somatic diversification of normal $CD5^-$ B cells appears to be initiated by antigen, require T cell help, and occur in germinal centers, whereas $CD5^+$ B cells are considered to be triggered by T cell–independent antigens, presumably in the mantle zones of lymph nodes (for review

see reference 48). However, although studies in humans support this general view (49), others also indicate that human $CD5^+$ B cells, both normal and leukemic, can respond to T cell–derived stimuli (50, 51). Indeed, the preferential use by our patients' $CD5^+$ B cells of $\gamma 1$ and $\gamma 3$ subclass genes and the absence of $\gamma 2$ (Table 1; 9) might favor the latter mechanism, since studies suggest that, in general, T cell–dependent protein antigens induce switching to IgG1 and IgG3, while switching to IgG2 frequently can be a function of T cell–independent polysaccharide antigens (for a review see reference 52). Finally, since at least in the murine system it is thought that $CD5^+$ B cells may represent a different cellular lineage (53) that follows a distinct differentiation pathway, such cells might use different mechanisms or anatomic sites to achieve somatic diversification. Similarly, diversification might occur in a distinct subset of these cells, due either to inherent subset differences or differences in accessory cell or T cell helper function. The latter might be especially relevant since we have shown previously that most IgM-producing CLL patients have diminished T cell helper function (54). The surface phenotypes of these IgG^+CD5^+ CLL B cells and their T cell functions are being evaluated to determine whether unique features exist for these particular cases.

The authors thank Dr. Harry W. Schroeder, Jr., for his advice and help in the analyses of the V_H gene sequences.

These studies were supported in part by U.S. Public Health Service grant AI-10811 from the National Institutes of Health National Institute of Allergy and Infectious Diseases, the Richard and Nancy Leeds Fund of the Department of Medicine of North Shore University Hospital, the Hematologic Research Foundation, the Joseph Eletto Leukemia Research Fund, and Consiglio Nazionale delle Richerche–Applicazione Cliniche della Ricerca Oncologica.

Address correspondence to Dr. Nicholas Chiorazzi, North Shore University Hospital, 350 Community Drive, Manhasset, NY 11030.

Received for publication 20 May 1994 and in revised form 21 December 1994.

References

1. Dighiero, G., P. Travade, S. Chevret, P. Fenaux, C. Chastang, and J.L. Binet. 1991. B-cell chronic lymphocytic leukemia: present status and future directions. *Blood.* 78:1901–1914.
2. Kipps, T. 1993. Immunoglobulin genes in chronic lymphocytic leukemia. *Blood Cells.* 19:615–625.
3. Broker, B.M., A. Klajman, P. Youinou, J. Jouquan, C.P. Worman, J. Murphy, L. Mackenzie, R. Quartey-Papafio, M. Blaschek, P. Collins, et al. 1988. Chronic lymphocytic leukemia cells secrete multispecific antibodies. *J. Autoimmun.* 1:469–481.
4. Sthoeger, Z.M., M. Wakai, D.B. Tse, V.P. Vinciguerra, S.L. Allen, D.R. Budman, S.M. Lichtman, P. Schulman, L.R. Weiselberg, and N. Chiorazzi. 1989. Production of autoantibodies by CD5-expressing B lymphocytes from patients with chronic lymphocytic leukemia. *J. Exp. Med.* 169:255–268.
5. Borche, L., A. Lim, J.L. Binet, and G. Dighiero. 1990. Evidence that chronic lymphocytic leukemia B lymphocytes are frequently committed to production of natural autoantibodies. *Blood.* 76:562–569.
6. Cai, J., C. Humphries, A. Richardson, and P.W. Tucker. 1992. Extensive and selective mutation of a rearranged V_H5 gene in human B cell chronic lymphocytic leukemia. *J. Exp. Med.* 176:1073–1081.
7. Rassenti, L.Z., and T.J. Kipps. 1993. Lack of extensive mutations in the V_H5 genes used in common B cell chronic lymphocytic leukemia. *J. Exp. Med.* 177:1039–1046.
8. Gearhardt, P.J., N.D. Johnson, R. Douglas, and L. Hood. 1981. IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. *Nature (Lond.).* 291:29–34.
9. Wakai, M., S. Hashimoto, M. Omata, Z.M. Sthoeger, S.L. Allen, S.M. Lichtman, P. Schulman, V.P. Vinciguerra, B. Diamond, M. Dono, et al. 1994. IgG^+, $CD5^+$ human chronic lymphocytic leukemia B cells. Production of IgG antibodies that exhibit diminished autoreactivity and IgG subclass skewing. *Autoimmunity.* 19:39–48.
10. Rai, K.R., A. Sawitsky, E.P. Cronkite, A.D. Chanana, R.N. Levy, and B.S. Pasternack. 1975. Clinical staging of chronic lymphocytic leukemia. *Blood.* 46:219–234.
11. Deane, M., and J.D. Norton. 1991. Immunoglobulin gene "fingerprinting": an approach to analysis of B lymphocyte clonality in lymphoproliferative disorders. *Br. J. Haematol.* 77:274–281.
12. Hingorani, R., I.H. Choi, P. Akolkar, B. Gulwani-Akolkar, R. Pergolizzi, J. Silver, and P.K. Gregersen. 1993. Clonal predominance of T cell receptors within the $CD8^+$ $CD45RO^+$ subset in normal human subjects. *J. Immunol.* 151:5762–5769.
13. Larrick, J.W., L. Danielsson, C.A. Brenner, E.F. Wallace, M. Abrahamson, K.E. Fry, and C.A.K. Borrebaeck. 1989. Polymerase chain reaction using mixed primers: cloning of human monoclonal antibody variable region genes from single hybridoma cells. *Biotechnology.* 7:934–938.
14. Harindranath, N., H. Ikematsu, A.L. Notkins, and P. Casali. 1993. Structure of the V_H and V_L segments of polyreactive and and monoreactive human natural antibodies to HIV-1 and *E. coli* B-galactosidase. *Int. Immunol.* 5:1523–1533.
15. Hashimoto, S., P.K. Gregersen, and N. Chiorazzi. 1993. The human Ig-β cDNA sequence, a homologue of murine B29, is identical in B cell and plasma cell lines producing all the human Ig isotypes. *J. Immunol.* 150:491–498.
16. Sanger, F., S. Nicklen, and A.R. Coulsen. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.
17. Elgavish, R., and H.W. Schroeder, Jr. 1993. SAW: a graphical user interface for the analysis of immunoglobulin variable domain sequences. *Biotechniques.* 15:1066–1071.
18. Willems van Dijk, K., E.H. Sasso, and E.C.B. Milner. 1991. Polymorphism of the human Ig V_H4 gene family. *J. Immunol.* 146:3646–3651.
19. Yamada, M., R. Wasserman, B.A. Reichard, S. Shane, A.H.

Caton, and G. Rovera. 1991. Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult peripheral blood B lymphocytes. *J. Exp. Med.* 173: 395–407.

20. Mayer, R., T. Logtenberg, J. Strauchen, A. Dimitriu-Bona, L. Mayer, S. Mechanic, N. Chiorazzi, L. Borche, G. Dighiero, A. Mannheimer-Lory, et al. 1990. CD5 and immunoglobulin V gene expression in B-cell lymphomas and chronic lymphocytic leukemia. *Blood.* 75:1518–1524.
21. Schroeder, H.W., and G. Dighiero. 1994. The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. *Immunol. Today.* 15:288–294.
22. Hashimoto, S., M. Wakai, J. Silver, and N. Chiorazzi. 1992. Biased usage of variable and constant region genes by IgG^+, $CD5^+$ human leukemic B cells. *Ann. NY Acad. Sci.* 651: 477–479.
23. Stevenson, F.K., M.B. Spellerberg, J. Treasure, C.J. Chapman, L.E. Silberstein, T.J. Hamblin, and D.B. Jones. 1993. Differential usage of an Ig heavy chain variable region gene by human B-cell tumors. *Blood.* 82:224–230.
24. Ebeling, S.B., M.E.M. Schutte, K.E. Akkermans-Koolhaas, A.C. Bloem, F.H.J. Gmelig-Meyling, and T. Logtenberg. 1992. Expression of members of the Ig V_H3 gene families is not restricted at the level of individual genes in human chronic lymphocytic leukemia. *Int. Immunol.* 4:313–320.
25. Kipps, T.J., S. Fong, E. Tomhave, P.P. Chen, R.D. Goldfien, and D.A. Carson. 1987. High frequency expression of a conserved κ light chain variable region gene in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA.* 84:2916–2920.
26. Mayumi, M., T. Kuritani, H. Kubagawa, and M.D. Cooper. 1983. IgG subclass expression by human B lymphocytes and plasma cells: B lymphocytes precommitted to IgG subclass can be preferentially induced by polyclonal mitogens with T cell help. *J. Immunol.* 130:671–677.
27. Froland, S.S., and J.B. Natvig. 1972. Class, subclass and allelic exclusion of membrane-bound Ig on human B lymphocytes. *J. Exp. Med.* 136:409–414.
28. Wortis, H.H. 1992. Surface markers, heavy chain sequences and B cell lineages. *Int. Rev. Immunol.* 8:235–246.
29. Haughton, G., L.W. Arnold, A.C. Whitmore, and S.H. Clarke. 1993. B-1 cells are made, not born. *Immunol. Today.* 14:84–87.
30. Herzenberg, L.A., A.M. Stall, P.A. Lalor, C. Sidman, W.A. Moore, D.R. Parks, and L.A. Herzenberg. 1986. The Ly-1 B cell lineage. *Immunol. Rev.* 93:81–102.
31. Victor, K.D., and J.D. Capra. 1994. An apparently common mechanism of generating antibody diversity: length variation of the V_L-J_L junction. *Mol. Immunol.* 31:39–46.
32. Lee, S.K., S.L. Bridges, Jr., W.J. Koopman, and H.W. Schroeder. 1992. The immunoglobulin kappa light chain repertoire expressed in the synovium of a patient with rheumatoid arthritis. *Arthritis Rheum.* 35:905–913.
33. Davidson, A., A. Smith, J. Katz, J.L. Preud'homme, A. Solomon, and B. Diamond. 1989. A cross-reactive idiotype on DNA antibodies defines a heavy chain determinant present almost exclusively on IgG antibodies. *J. Immunol.* 143:174–180.
34. Davidson, A., J.L. Preud'homme, A. Solomon, M.-D. Chang, S. Beede, and B. Diamond. 1987. Idiotypic analysis of myeloma proteins: anti-DNA activity of monoclonal immunoglobulins bearing an SLE idiotype is more common in IgG than IgM antibodies. *J. Immunol.* 138:1515–1518.
35. Pritsch, O., C. Magnac, G. Dumas, C. Egile, and G. Dighiero. 1993. V gene usage by seven hybrids derived from $CD5^+$ B-cell chronic lymphocytic leukemia and displaying autoantibody activity. *Blood.* 82:3103–3112.
36. Roudier, J., G.J. Silverman, P.P. Chen, D.A. Carson, and T.J. Kipps. 1990. Intraclonal diversity in the V_H genes expressed by CD5-negative chronic lymphocytic leukemia producing pathogenic IgM rheumatoid factor. *J. Immunol.* 144:1526–1529.
37. Wagner, S.D., and L. Luzzatto. 1993. V_κ gene segments rearranged in chronic lymphocytic leukemia are distributed over a large portion of the V_κ locus and do not show somatic mutation. *Eur. J. Immunol.* 23:391–397.
38. Mantovani, L., R.L. Wilder, and P. Casali. 1993. Human rheumatoid B-1a ($CD5^+$ B) cells make somatically hypermutated high affinity IgM rheumatoid factors. *J. Immunol.* 151:473–488.
39. Jukes, T.H., and J.L. King. 1979. Evolutionary nucleotide replacements in DNA. *Nature (Lond.).* 281:605–606.
40. Radic, M.Z., M.A. Mascelli, J. Erikson, H. Shan, M. Shlomchik, and M. Weigert. 1989. Structural patterns in anti-DNA antibodies from MRL/lpr mice. *Cold Spring Harbor Symp. Quant. Biol.* 54:933–1001.
41. Radic, M.Z., J. Mackle, J. Erikson, C. Mol, W.F. Anderson, and M. Weigert. 1993. Residues that mediate DNA binding of autoimmune antibodies. *J. Immunol.* 150:4966–4977.
42. Bhat, T.N., G.A. Bentley, T.O. Fischmann, G. Boulot, and R.J. Poljak. 1990. Small rearrangements in structures of Fv and Fab fragments of antibody D1.3 on antigen binding. *Nature (Lond.).* 347:483–485.
43. Friedman, D.F., E.A. Cho, J. Goldman, C.E. Carmack, E.C. Besa, R.R. Hardy, and L.E. Silberstein. 1991. The role of clonal selection in the pathogenesis of an autoreactive human B cell lymphoma. *J. Exp. Med.* 174:525–537.
44. Bahler, D.W., and R. Levy. 1992. Clonal evolution of a follicular lymphoma: evidence for antigen selection. *Proc. Natl. Acad. Sci. USA.* 89:6770–6774.
45. Jain, R., S. Roncella, S. Hashimoto, A. Carbone, P. Francia di Celle, R. Foa, M. Ferrarini, and N. Chiorazzi. 1994. A potential role for antigen selection in the clonal evolution of Burkitt's lymphoma. *J. Immunol.* 153:45–52.
46. Diamond, B., and M.D. Scharff. 1984. Somatic mutations of the T15 heavy chain gives rise to an antibody with autoantibody specificity. *Proc. Natl. Acad. Sci. USA.* 81:5841–5844.
47. Chen, C., V.A. Roberts, and M.B. Rittenberg. 1992. Generation and analysis of random point mutations in an antibody CDR2 sequence: many mutated antibodies lose their ability to bind antigen. *J. Exp. Med.* 176:855–866.
48. MacLennan, I.C.M. 1994. Germinal centers. *Annu. Rev. Immunol.* 12:117–139.
49. Zupo, S., M. Dono, L. Azzoni, N. Chiorazzi, and M. Ferrarini. 1991. Evidence for differential responsiveness of human $CD5^+$ and $CD5^-$ B cell subsets to T cell-independent mitogens. *Eur. J. Immunol.* 21:351–359.
50. Fu, S.M., N. Chiorazzi, H.G. Kunkel, J.P. Halper, and S.R. Harris. 1978. Induction of in vitro differentiation and immunoglobulin synthesis of human leukemic B lymphocytes. *J. Exp. Med.* 148:1570–1578.
51. Zupo, S., M. Dono, R. Massara, G. Taborelli, N. Chiorazzi, and M. Ferrarini. 1994. Expression of CD5 and CD38 by human $CD5^-$ B cells: requirement for special stimuli. *Eur. J. Immunol.* 24:1426–1433.
52. Esser, C., and A. Radbruch. 1990. Immunoglobulin class switching: molecular and cellular analysis. *Annu. Rev. Immunol.* 8:717–735.
53. Hayakawa, K., R.R. Hardy, and L.A. Herzenberg. 1985. Progenitors for Ly-1 B cells are distinct from the progenitors of other B cells. *J. Exp. Med.* 161:1554–1568.
54. Chiorazzi, N., S.M. Fu, G. Montazeri, H.G. Kunkel, J.P. Halper, and S. Harris. 1979. T cell helper defect in patients with chronic lymphocytic leukemia. *J. Immunol.* 122:1087–1090.