

# Human heavy-chain variable region gene family nonrandomly rearranged in familial chronic lymphocytic leukemia

(DNA rearrangements/B lymphocytes/somatic mutation)

A. SHEN\*†, C. HUMPHRIES‡, P. TUCKER‡, AND F. BLATTNER\*§

†Department of Microbiology, University of Texas Health Science Center, Dallas, TX 53223; and \*Department of Genetics, University of Wisconsin, Madison, WI 53706

Communicated by Elvin A. Kabat, June 29, 1987

**ABSTRACT** We have identified a family of human immunoglobulin heavy-chain variable-region ( $V_H$ ) genes, one member of which is rearranged in two affected members of a family in which the father and four of five siblings developed chronic lymphocytic leukemia. Cloning and sequencing of the rearranged  $V_H$  genes from leukemic lymphocytes of three affected siblings showed that two siblings had rearranged  $V_H$  genes ( $V_{HTS1}$  and  $V_{HWS1}$ ) that were 90% homologous. The corresponding germ-line gene,  $V_{H251}$ , was found to be part of a small (four gene)  $V_H$  gene family, which we term  $V_{HV}$ . The DNA sequence homology to  $V_{HWS1}$  (95%) and  $V_{HTS1}$  (88%) and identical restriction sites on the 5' side of  $V_H$  confirm that rearrangement of  $V_{H251}$  followed by somatic mutation produced the identical  $V_H$  gene rearrangements in the two siblings.  $V_{HTS1}$  is not a functional  $V_H$  gene; a functional  $V_H$  rearrangement was found on the other chromosome of this patient. The other two siblings had different  $V_H$  gene rearrangements. All used different diversity genes. Mechanisms proposed for non-random selection of a single  $V_H$  gene include developmental regulation of this  $V_H$  gene rearrangement or selection of a subpopulation of B cells in which this  $V_H$  has been rearranged.

The prevalent form of adult leukemia in western countries is chronic lymphocytic leukemia (CLL), in which the normally diverse population of B lymphocytes is replaced by a nearly monoclonal accumulation of cells (1). The disease has a tendency to cluster in families (2, 3) indicating that genetic factors may be involved in its etiology.

It is not yet known whether particular gene arrangements turn up in CLL, especially in familial cases. Variable-region gene selection could be purely random or quite biased if the disease were to directly influence DNA rearrangement or antigen driven clonal expansion. The study of CLL variable regions at the DNA sequence level may shed light on the genetic and/or environmental factors involved in the disease. We have analyzed three members of a family in which the father and four of five siblings developed CLL (4). This family has been studied by W. Blattner and coworkers (5) who showed that the cells express surface IgM and IgD with  $\kappa$  light chains. We have found rearrangement of the same heavy-chain variable-region ( $V_H$ ) gene in the two male siblings of this family. This particular  $V_H$  gene has an unusual DNA sequence that we suggest defines a previously unrecognized  $V_H$  gene family. It is represented in the human genome by, at most, four homologous sequences.

## MATERIALS AND METHODS

Frozen CLL lymphocytes obtained by leukaphoresis and fibroblasts of the affected siblings were provided by W.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Blattner (Environmental Epidemiology Branch, National Cancer Institute, Frederick, MD).

DNA was isolated from patient lymphocytes or fibroblasts, digested with restriction enzymes, electrophoresed on 1% agarose gels (10  $\mu$ g of DNA per lane), blotted onto nitrocellulose (6), and hybridized to  $^{32}$ P-labeled probes prepared as described (7). Hybridization was carried out at 68°C in a solution of 1 M NaCl, 50 mM sodium phosphate (pH 6.5), 2 mM EDTA, 10 $\times$  Denhardt's solution, denatured, sonicated salmon sperm DNA at 100  $\mu$ g/ml, and 0.5% NaDodSO<sub>4</sub>. Filters were washed at 68°C with 1 $\times$  SSC/0.1% NaDodSO<sub>4</sub> ( $V_H$  probes) or 0.1 $\times$  SSC/0.1% NaDodSO<sub>4</sub> ( $J_H$  probes). (1 $\times$  Denhardt's solution = 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin. 1  $\times$  SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0.)

Partial *Sau3A* libraries in Charon 30 (8) or complete *HindIII* or *EcoRI* libraries in Charon 35 (9) were constructed from patient lymphocyte DNA (6). The unamplified libraries were screened with the probe p $J_H$  (see Fig. 2). A Charon 35 *HindIII* library prepared from leukocyte DNA from an unrelated individual was screened with a 187-base-pair (bp) *Bgl* II-*Pst* I fragment covering  $V_{HWS1}$  to isolate the germ-line  $V_H$  gene.

Restriction fragments hybridizing to the immunoglobulin heavy-chain joining-region genes ( $J_H$ ) were subcloned into pUC13 (10) and sequenced by the method of Maxam and Gilbert (11).

## RESULTS

**Common  $J_H$ -Region Rearrangements in Leukemic Lymphocytes of CLL Family Members.** Immunoglobulin gene rearrangements in family members were identified using a human  $J_H$  gene probe. This probe covering  $J_{H2}$  to  $J_{H6}$  (see J2J6 in Fig. 2A) can identify  $J_H$  rearrangements since the  $V_H$ - $D_H$ - $J_H$  (where  $D_H$  is the immunoglobulin heavy-chain diversity-region gene) joining leads to deletion of the intervening DNA and altered restriction enzyme sites on the 5' side of the  $J_H$  region (12). Fig. 1A shows the family pedigree and pattern of J2J6 hybridization to fibroblast (germ-line) and lymphocyte (rearranged) DNA. Fibroblast DNAs digested with *HindIII* and *Bam*HI had a single 6.0-kbp J2J6-hybridizing band that is the expected germ-line size (14). Lymphocytes from patients HP, TS, and WS had two rearranged  $J_H$ -hybridizing bands characteristic of a monoclonal population plus various levels of the unrearranged 6.0-kbp band (presumably due to the presence of nonleukemic cells). TS and WS have common rearranged bands of  $\approx$ 8.5 kbp and 2.8 kbp, whereas HP has 4.5-kbp and 3.0-kbp rearranged bands. DNA from PR, who

Abbreviations: CLL, chronic lymphocytic leukemia;  $V_H$ ,  $D_H$ , and  $J_H$ , heavy-chain variable-, diversity-, and joining-region genes; CDR, complementarity determining region; FR, framework region.

†Present address: Department of Oncology, University of Wisconsin, Madison, WI 53706.

§To whom reprint requests should be addressed.

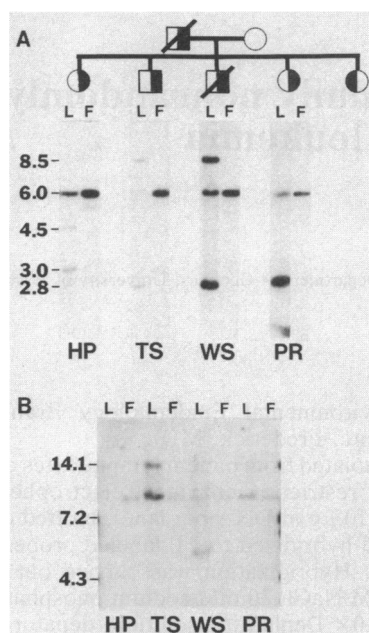


FIG. 1. Family pedigree and analysis of immunoglobulin  $J_H$  and  $J_K$  rearrangements.  $\blacksquare$ , Male with CLL;  $\bullet$ , female with CLL;  $\circ$ , female with other neoplasm;  $\diagup$ , deceased. (A) Immunoglobulin  $J_H$  rearrangements. Lymphocyte (lanes L) or fibroblast (lanes F) DNA was digested with *Hind*III and probed with J2J6, a 2.4 kbp *Sau*3A fragment carrying  $J_H2$ – $J_H6$  (see Fig. 2). Fragment sizes in kbp are indicated at the left, and patients are indicated at the bottom. (B) Immunoglobulin  $J_K$  rearrangements. Lymphocyte (lanes L) or fibroblast (lanes F) DNA was digested with *Hind*III and probed with a 1.9-kbp *Sac* I fragment covering  $J_K$  (13). Fragment sizes in kbp are indicated at the left and patients are indicated at the bottom.

is in spontaneous remission, showed a germ-line band and a single 2.9-kbp rearranged band. *Eco*RI digests also showed that TS and WS shared identical rearrangements (data not shown).

Since all the siblings expressed surface  $\kappa$  light chain, a similar analysis was carried out with a  $\kappa$  light-chain probe (13). Fig. 1B shows no common light-chain rearrangements in this family.

To determine if the common  $J_H$ -hybridizing bands were, in fact, due to identical  $V_H$ - $D_H$ - $J_H$  rearrangements, the rearranged genes were cloned and the  $V_H$ - $D_H$ - $J_H$  rearrangements were sequenced.

The 8.5- and 2.8-kbp  $J_H$ -hybridizing bands seen in Fig. 1A were cloned in separate phages and represent the two rearranged chromosomes (designated 1 and 2) of TS and WS. The 3.0-kbp  $J_H$ -hybridizing band of HP corresponds to the phage clones designated HP1. Fig. 2A shows the restriction maps of the common 8.5-kbp TS1 and WS1 rearrangements together with the corresponding germ-line  $J_H$  and  $V_H$  restriction maps whereas Fig. 2B shows the other rearrangements.

**A Single Germ-Line Gene from a Previously Unidentified Low-Copy-Number Human  $V_H$  Gene Family is Rearranged in Both TS and WS.** Probes specific for the  $V_H$ -coding region or the 5'-flanking region of WS1 and TS1 were hybridized to *Hind*III digests or *Hind*III/*Bam*HI double digests of patient lymphocyte or fibroblast DNA or of normal liver DNA. No differences in hybridization patterns were seen with the various probes. All hybridized strongly to the rearranged 8.5-kbp *Hind*III-*Bam*HI fragment in TS and WS lymphocytes (data not shown) and to 7.5-, 3.2-, 1.5-, and 1.4-kbp germ-line *Hind*III-*Bam*HI bands (Fig. 3A). *Hind*III digests showed two strongly hybridizing germ-line bands (10.5 and 5.5 kbp). (Fig. 3B); the 10.5-kbp band in TS was replaced by a 11.2-kbp rearranged band. The 10.5- and 5.5-kbp *Hind*III

bands were not polymorphic among unrelated DNAs. Depending on the enzyme used, only three or four bands were ever detected, indicating that  $V_H$ TS1 and  $V_H$ WS1 were derived from a small family of  $V_H$  genes.

Germ-line clones representing the four *Hind*III-*Bam*HI bands of Fig. 3A were isolated and sequenced from DNA of unrelated individuals. The phage clones corresponding to the most strongly hybridizing 7.5-kbp *Hind*III-*Bam*HI fragment are shown in Fig. 2A; the  $V_H$  gene is called  $V_H$ 251. The phages carrying  $V_H$ 251,  $V_H$ WS1, and  $V_H$ TS1 have nearly identical restriction maps on the 5' side of the  $V_H$  region, suggesting that  $V_H$ WS1 and  $V_H$ TS1 arose from rearrangement of  $V_H$ 251.

The comparison of the DNA and amino acid sequences of the WS1 and TS1  $V_H$ - $D_H$ - $J_H$  genes with that of the germ-line  $V_H$ 251 and  $J_H$ 4 genes is shown in Fig. 4. Both TS1 and WS1 used  $J_H$ 4 and the same  $V_H$  segments but different  $D_H$  segments.  $V_H$ WS1 was productively rearranged. Both  $V_H$ WS1 and the germ-line  $V_H$ 251 had all the features characteristic of a functional  $V_H$  gene—a conserved octanucleotide sequence ATGCAAT located 105 bp before the translation initiation codon (16), a leader sequence with appropriate junctions for splicing to the  $V_H$  (17), and conserved residues (e.g., cysteines at positions 22 and 92) at anticipated positions (15).  $V_H$ TS1 was 90% homologous to  $V_H$ WS1 but cannot be expressed. It has terminators at amino acids 46 and 83, a thymidine to adenosine base change that eliminates the translation initiation codon at position -19, and a frame shift at the  $D_H$ - $J_H$  junction that alters the  $J_H$ 4 reading frame.

The DNA sequence homology of the germ-line  $V_H$ 251 gene to  $V_H$ WS1 (95%) and  $V_H$ TS1 (88%) and identical 5'-restriction sites confirm that  $V_H$ 251 was indeed the precursor to both  $V_H$ WS1 and  $V_H$ TS1. The other germ-line  $V_H$  genes are less homologous (82% and 71%, respectively, with  $V_H$ WS1) and the 5'-restriction sites do not align with those of the  $V_H$ TS1 or  $V_H$ WS1 genes (data not shown).

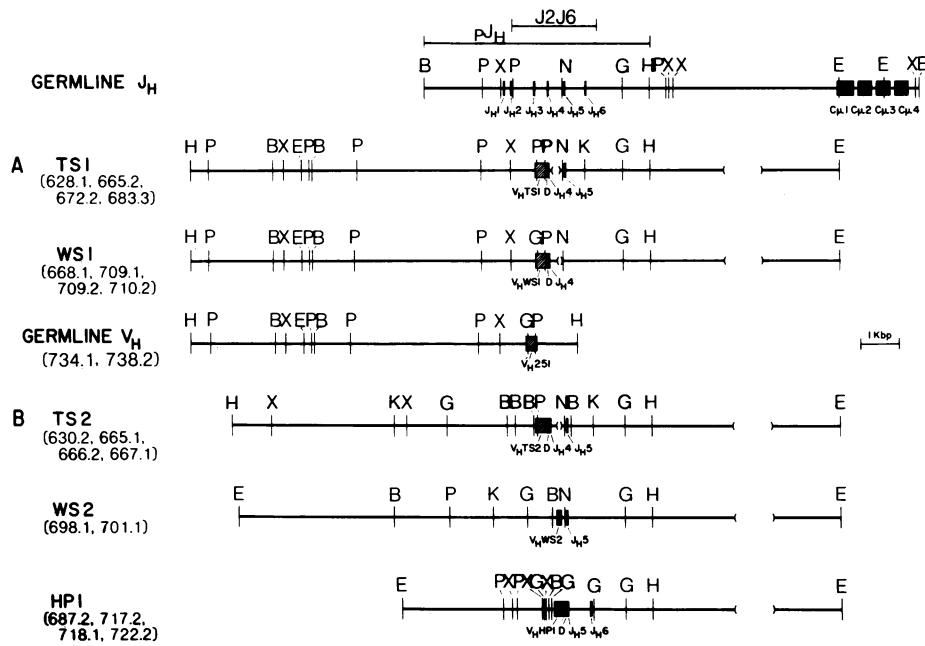
$V_H$ WS1 differs by 21 bp from the germ-line gene. Substitutions were most frequent in complementarity determining regions (CDRs), with 3 base changes in CDR1 and 4 in CDR2. The proportion of replacement substitutions (those producing an amino acid change) ranges from 0.3 to 0.5 in the framework region (FR); in the CDRs the replacement proportion is 1.0.  $V_H$ TS1 had the most amino acid changes in CDR1 (4/5), CDR2 (4/17), and FR3 (10/32). The proportion of replacement substitutions in the CDRs was 1.0 and ranged from 0.5 in FR2 to 1 in FR3. A few of the observed changes, especially those in which  $V_H$ TS1 and  $V_H$ WS1 were the same but different from the germ line, could reflect the fact that the germ-line DNA came from an unrelated individual.

The germ-line sequence appears to constitute a previously unidentified  $V_H$  gene family that we propose to designate  $V_H$ V. Although its 83-bp intron size is consistent with that of the  $V_H$ I family, its closest relatives in the GenBank,<sup>†</sup> HG3 (18) and LR-35, are only 64% and 62% homologous. The two  $D_H$  segments are also different from any reported.

**Analysis of the Other  $J_H$ -Region Rearrangements.** The other chromosome of TS, TS2, carried a productively rearranged  $V_H$  gene ( $V_H$ TS2) (Fig. 5). This was a  $V_H$ II gene by its 82% DNA sequence homology to another human  $V_H$ II gene (20) and amino acid homology to other human  $V_H$ II proteins, WAH (21) and Ou (22). The second chromosome of WS, WS2, carried an aberrant rearrangement in which the 3' end of the  $V_H$  gene and the  $D_H$  and  $J_H$ 4 segments had been deleted and the coding sequence for the first 40 amino acids of the  $V_H$  gene was joined to the  $J_H$ 4- $J_H$ 5 intron (data not shown).

Recall that TS and WS also shared a commonly migrating 2.8-kbp  $J_H$ -hybridizing band (Fig. 1A). Sequence analysis

<sup>†</sup>EMBL/GenBank Genetic Sequence Database (1986) GenBank (Bolt, Beranek, and Newman Laboratories, Cambridge, MA), Tape Release 42.0.



**FIG. 2.** Restriction maps of cloned  $V_H$ - $D_H$ - $J_H$  rearranged genes and corresponding germ-line  $V_H251$  gene.  $\square$ ,  $V_H$  genes;  $\square$ ,  $D_H$  genes;  $\blacksquare$ ,  $J_H$  and  $\mu$ -chain constant-region ( $C_\mu$ ) genes. Parentheses indicate deletions relative to the published  $J_H$ - $C_\mu$  restriction map (14). The germ-line  $J_H$  map (14) and restriction fragments used for probes are shown above the rearranged genes. Phage clones covering the rearranged genes are indicated at the left. E, *EcoRI*; H, *HindIII*; B, *BamHI*; P, *Pst I*; X, *Xba I*; G, *Bgl II*; K, *Kpn I*; N, *Nar I*. (A) Restriction maps of rearranged TS1 and WS1 chromosomes carrying the common 8.5-kbp *HindIII*-*BamHI*  $J_H$  rearrangement and corresponding germ-line  $V_H251$  gene. (B) Restriction maps of the other cloned  $J_H$  rearrangements. TS2 and WS2 carry the 2.8-kbp *HindIII*-*BamHI* fragments of TS and WS, respectively. HP1 carries the 3.0-kbp *HindIII*-*BamHI* fragment of patient HP.

revealed that this was not due to a common rearrangement, but to the following two separate events: a mutation in the  $J_H4$ - $J_H5$  intron creating a *BamHI* site in TS2 and a rearrangement coupled with deletions in WS2.

Sequencing of HP1 phage clones, carrying the 3.0-kbp *HindIII*-*BamHI*  $J_H$ -hybridizing band, showed a productively rearranged  $V_H$ - $D_H$ - $J_H5$  (Fig. 5). The DNA sequence of  $V_H$ HP1 was 80% homologous to the  $V_H I$  genes HA2 and HG3 (18).

We did not clone the 4.5-kbp *HindIII*-*BamHI*  $J_H$ -hybridizing band of HP.  $V_H$ WS1 did not cross-hybridize to this 4.5-kbp band. Hybridization of  $V_H251$  to HP lymphocyte DNA showed only germ-line bands, indicating that  $V_H251$  was not rearranged to any detectable extent in HP (Fig. 3B). Hybridization of  $V_H$ WS1 to lymphocyte DNA of PR also showed no detectable rearrangement (data not shown). The 2.9-kbp rearranged band of PR was not further characterized.

**Base Changes and Deletions in the  $J_H4$ - $J_H5$  Intron Accompany Rearrangement.** The DNA sequences of  $J_H4$  and the  $J_H4$ - $J_H5$  intron in TS1, WS1, and TS2 vary considerably from the published germ-line sequence of this region (14). The  $J_H4$ - $J_H5$  introns of TS1 and WS1 contain significant deletions of 280 bp and 70 bp, respectively. Two independently isolated TS1 phages had identical deletions in this region. The deletion in TS1 leading to a 8.2-kbp, rather than a 8.5-kbp,  $J_H$ -hybridizing band has been confirmed by a genomic Southern blot (data not shown). The *Kpn I* sites in TS2 and TS1 and *BamHI* site in TS2 were found in lymphocyte DNA only, not in fibroblasts from TS. These differences do not appear to result from polymorphism, but more probably reflect somatic mutation, perhaps during the rearrangement process.

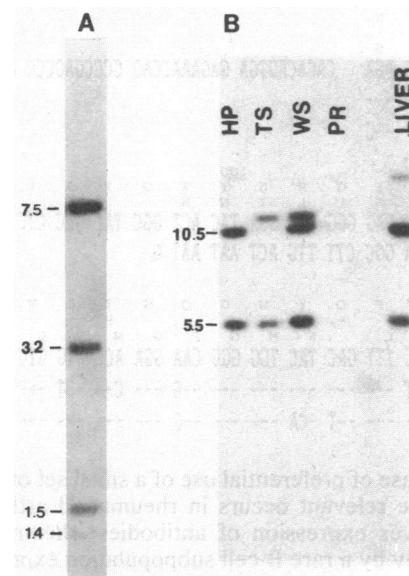
We did not find evidence for translocations into the  $J_H$  region as has been reported in two other CLL cases (23, 24). A region some 10 kbp on the 5' side of WS1 did hybridize to multiple bands on a genomic Southern blot (data not shown); however, preliminary analysis suggests it was another  $V_H$  gene.

### DISCUSSION

We have isolated and sequenced the immunoglobulin heavy-chain  $V_H$ - $D_H$ - $J_H$  gene rearrangements in three siblings with CLL and found that two siblings, TS and WS, had rearranged the same germ-line  $V_H$  gene ( $V_H251$ ) in their leukemic lymphocytes, the former nonproductively and the latter productively.  $V_H251$  was shown to be a single-copy gene

belonging to a small, previously unrecognized, gene family we term  $V_H V$ . A third sibling, HP, rearranged a different  $V_H$  gene. All the rearrangements used different  $D_H$  segments. Random rearrangement to a single-copy  $V_H V$  gene in two individuals should be rare (frequency < 0.003, assuming the total number of human  $V_H$  genes is 200; ref. 19 and J. Berman and F. Alt, personal communication).

Preferential  $V_H$ -gene rearrangement occurs in mice where the  $V_H7183$  family, inferred to be closest to the  $D_H$  cluster, is frequently used in the earliest B cells (25, 26). Most clones of this type are eliminated subsequently from the B-cell population. The biased rearrangement of  $V_H251$  may be explained on a similar basis if this  $V_H V$  gene is close to  $D_H$  on the human chromosome, since CLL cells represent an early stage of B-cell development.



**FIG. 3.** Southern blot analysis of germ-line  $V_H251$  genes. Lymphocyte DNA from the four affected siblings (indicated by initials) or normal liver DNA was digested and probed with a 1-kbp *Stu I*-*Xba I* fragment carrying the  $V_H251$  gene and its 5'-flanking region. Fragment sizes in kbp are indicated. (A) Normal liver DNA digested with *HindIII* and *BamHI*. (B) Patient lymphocyte DNA and normal liver DNA digested with *HindIII*.

V <sub>H</sub> 251	ATGCAATGC A-GTGGGGC CTCCCCACTT AAACCCAGGG CTCCCCTCCA CAGTGAGTCT CCCTCACTGC CCAGCTGGGA TCTCAGGGCT
V <sub>H</sub> WS1	----- -A----- -G----- T-----
V <sub>H</sub> TS1	----- -A----- *----- T-----
V <sub>H</sub> 251	<sup>-19</sup> M G S T A I L A L L L A I L Q
V <sub>H</sub> WS1	- - - - -
V <sub>H</sub> TS1	- - - - -
V <sub>H</sub> 251	TCATTTTCTG TCCTCCACCATC ATG GGG TCA ACC GCC ATC CTC GCC CTC CTC CTA GCT ATT CTC CAA G/ GTCAGTCTT
V <sub>H</sub> WS1	----- -G----- G-----
V <sub>H</sub> TS1	----- -A----- -G----- GC-----
V <sub>H</sub> 251	<sup>-4</sup> G V C A
V <sub>H</sub> WS1	- - - - -
V <sub>H</sub> TS1	- - - - -
V <sub>H</sub> 251	GCCGAGGGCT TGAGGTCA GAGGAGAACC G-TGGAAAGG AGCCCCTGAT TCAAATTTG TGTCTCCCC ACAG /GA GTC TGT GCC
V <sub>H</sub> WS1	----- -G-----
V <sub>H</sub> TS1	----- -G-----
V <sub>H</sub> 251	<sup>1</sup> E V Q L V Q S G A <sup>10</sup> E V K K P G E S L K <sup>20</sup> I S C K G
V <sub>H</sub> WS1	- - - - -
V <sub>H</sub> TS1	- - - - -
V <sub>H</sub> 251	GAG GTG CAG CTG GTG CAG TCT GGA GCA GAG GTG AAA AAG CCC GGG GAG TCT CTG AAG ATC TCC TGT AAG GGT
V <sub>H</sub> WS1	----- -A----- -G-----
V <sub>H</sub> TS1	--- C--- -G----- -G-----
V <sub>H</sub> 251	<sup>30</sup> S G Y S F T S Y W T G W V R Q M P G K G L E W M
V <sub>H</sub> WS1	- - - - -
V <sub>H</sub> TS1	- - - - -
V <sub>H</sub> 251	TCT GGA TAC AGC TTT ACC AGC TAC TGG ACC GGC TGG GTG CGC CAG ATG CCC GGG AAA GGC TTG GAG TGG ATG
V <sub>H</sub> WS1	----- -G----- -C----- -T----- -C-----
V <sub>H</sub> TS1	-A--C- ---G--- -GT--G- ---T--C- ---T----- -G C-A T-- -TA---
V <sub>H</sub> 251	<sup>50</sup> G I I Y <sup>52A</sup> G D S D T R Y <sup>60</sup> S P S F Q G O V T I <sup>70</sup> S A
V <sub>H</sub> WS1	- - - - -
V <sub>H</sub> TS1	- - - - -
V <sub>H</sub> 251	GGG ATC ATC TAT CCT GGT GAC TCT GAT ACC AGA TAC AGC CCG TCC TTC CAA GGC CAG GTC ACC ATC TCA GCC
V <sub>H</sub> WS1	----- G----- -G----- -C----- -G-----
V <sub>H</sub> TS1	----- -C----- T----- CC----- A----- -G----- T-----
V <sub>H</sub> 251	<sup>80</sup> D K S I S T A Y L Q <sup>82</sup> W S <sup>82A</sup> S <sup>82B</sup> S <sup>82C</sup> L K A S D T A M <sup>90</sup> Y Y C
V <sub>H</sub> WS1	- - - - -
V <sub>H</sub> TS1	- - - - -
V <sub>H</sub> 251	GAC AAG TCC ATC AGC ACC GCC TAC CTG CAG TGG AGC AGC CTG AAG GCC TCG AC ACC GCC ATG TAT TAC TGT
V <sub>H</sub> WS1	----- -C----- -T-----
V <sub>H</sub> TS1	----- -G T----- -G----- T----- -G----- -G----- -A-----
V <sub>H</sub> 251	A R
V <sub>H</sub> WS1	- -
V <sub>H</sub> TS1	- S
V <sub>H</sub> 251	GCG AGA CACACAGTGA GAGAAACCAG CCCCGAGCCC GTCTAAAACC CTCCACACGT AGGGAGTAAT GAG
V <sub>H</sub> WS1	-----
V <sub>H</sub> TS1	---C-
D <sub>H</sub> WS1	<sup>100</sup> L E G R L T N G Y A L P Y
D <sub>H</sub> TS1	- - - - -
D <sub>H</sub> WS1	CTG GAG GGG CGT GGA TAC ACT GGC TAC GCC CTC CCC TAT
D <sub>H</sub> TS1	CGA GGC CTT TTG ACT AAT AAT G
J <sub>H</sub> 4(WS1)	<sup>110</sup> Y F D Y W G Q G T L V T V S S
J <sub>H</sub> 4(TS1)	- - - - -
J <sub>H</sub> 4	TAC TTT GAC TAC TGG GGC CAA GGA ACC CTG GTC ACC GTC TCC TCA G/GTAAGAATGGCTCT
J <sub>H</sub> 4(WS1)	--T----- -G----- C----- T-----
J <sub>H</sub> 4(TS1)	AC----- -T-CA----- -G----- -C-----

FIG. 4. Comparison of the DNA and amino acid sequences of WS1 and TS1  $V_H$ - $D_H$ - $J_H$  genes with the germ-line  $V_H$ 251 and  $J_H$ 4 genes. The DNA sequences begin with the conserved octamer sequences. The predicted amino acid sequences are shown above the DNA sequences and are placed at the center of each triplet. Numbering is according to Kabat *et al.* (15). The  $J_H$ 4 sequence is taken from Ravetch *et al.* (14). Dashes indicate identity with the germ-line sequence. Gaps have been placed at position 12 of the  $V_H$ 251 DNA sequence (-) and position 33 of the  $V_H$ TS1 DNA sequences (\*) to align the sequences. The insertion at the  $D_H$ - $J_H$  junction of the TS1 gene is shown, and the  $J_H$ 4 DNA sequence of TS1 is aligned with the germ-line  $J_H$ 4. Single-letter amino acid code is used. RNA splicing positions are indicated by a slash.

Another case of preferential use of a small set of  $V_H$  genes that might be relevant occurs in rheumatoid arthritis (27), which involves expression of antibodies (rheumatoid factors), possibly by a rare B-cell subpopulation expressing the antigen Leu 1 (28). CLL cells also express Leu 1 (1). The underlying basis of rheumatoid arthritis is, of course, unknown.

Other ways to account for nonrandom  $V_H$ 251 gene utilization involve selection by antigens, T-cells, or virus interacting with cell-surface  $V_H$ 251 protein. This is difficult to explain in the case of the nonproductively rearranged  $V_H$ 251 gene of TS unless the gene had at one time been expressed,

the cells selected, and the gene later inactivated by somatic mutation. In  $V_H$ WS1 the mutations were clustered in the CDRs with a high ratio of replacement to silent substitutions. In the FR, this ratio was low. This pattern was observed in the course of an antigen response (29) and indicates conservation of framework functions and selection for receptor function at the protein level. TS1 showed less clustering in CDRs and a very high proportion of replacements throughout, which is not consistent with receptor function.

In this connection, Cleary *et al.* (30) have reported rapid accumulation of somatic mutations in the  $V_H$  region of a human B-cell lymphoma during the course of therapeutic

**TS2**

```

CTCTTTTCACAGGGTCTGTCACAGGTCAGCTGCGAGGAGTGGGCGCCAGGACTGGTGAAGCCTTCACAGGCCCTGGCCCTCACCTGCACTGTCTCTG 100
/G V L S Q V Q L Q E S G P G L V K P S Q A L A L T C T V S
-4
GTGGCTCCTCAGCACTGATAAATACTACTGAGCTGGCTCCGGCAGCCCGCCGGGAAGGACTGGAGTGGATCGGGCGTATTTTGGCAGTGGCGCAC 200
G G S F S T D N Y Y W S W L R Q P A G K G L E W I G R I F G S G R T
CCACTTCAACCCGTCCTGGAAGAGTCGAGTCACCATATCAGTCGACACGTCCAA AAAACAGTTCTCCCTGATGTTGACCTCCGTGACCCCGCAGACTCG 300
H F N P S L K S R V T I S V D T S K N Q F S L M L T S V T A A D S
GCCGTATATTATTGTGCGAGAGGCTATGATCTGCGGGATATATGGCCCGCGCTTTGACTACTGGGGCCGGGAACCCGGTTACCGTCTCTCTCA 396
A V Y Y K A R G Y D T G G Y M A R G F D Y W G P G T R V T V S S
          92
    
```

**HP1**

```

ATGCAAAACACCTGAGATTTACTGAAGTAAATACAGATCTGCTGCTGCGCCCTGAGAGCATCACCCAGCAACCCATCTGCTCTAGAGAATCCCCTGAG 100
AGCTCCGTTCTCACCATGGACTGGACCTGGAGGATCCTCTTCTGGTGGCAGCAGCCAGTAAGAGGCTCCCTAGTCCAGTGATGAGAAAGAGATT 200
M D W T W R I L F L V A A T /
-19
GAGTCCAGTCCAGGAGATCTCATCCACTCTGTTCTCTCCACAGGAGCCACTCCAGGAGCAGCTGGTACAGTCTGGGGCTGAAGTGAAGAAGCCT 300
/G A H S Q E Q L V Q S G A E V K K P
-4
GGGCGCTCAGTGAAGTCTCATGCAAGGCTTGGATACAGCTTCATCGTATCTATATACACTGGCTCGCTCAGGCCCTGGACAAGGGCTTGAGTGA 400
G A S V K V S C K A S G Y S F I V Y Y I H W L R Q A P G Q L E W
TGGATGGATCAACCTAGGACCGGTGACACAACTATGCACAGACCTTTCAGGGCAGGGTCCACATGACCCAGGACAGCTCCATCAGTACGGCTACAT 500
M G W I N P R T G D T N Y A Q T F Q G R V T H T R D T S I S T A Y H
GGACCTGGGCGAGGCTGAGATTGACGACAGCCGCTTACTGTGCGAGAGATGTTGGAGTTGGGTACAGGCTGGTTCGACCCCTGGGGCCAG 600
D L G R L R F D D T A V Y Y C A R D V E L R Y G T G W F D P W H G Q
          92
GGAACCTGGTCACCGTACC 621
G T L V T V T
          DH JH5
    
```

FIG. 5. Sequences of HP1 and TS2 V<sub>H</sub>-D<sub>H</sub>-J<sub>H</sub> genes. The predicted amino acids are shown below the DNA sequence and are abbreviated as in Fig. 4. The octamer, D<sub>H</sub>, and J<sub>H</sub> sequences are underlined.

anti-idiotype treatment leading to elimination of the targeted V<sub>H</sub> idiotope, which they interpreted as selection against cells expressing the idiotope. Although TS and WS did not receive anti-idiotype therapy, the occurrence of mutations over a decade in response to the host defenses could follow a similar pattern.

A correlation may also be drawn between expression of V<sub>H</sub>251 and the severity of the disease. WS, with the expressed V<sub>H</sub>251 gene and an aberrant rearrangement on the other chromosome, had the most aggressive progression of disease; whereas TS, with an unexpressed V<sub>H</sub>251, showed a change to a more "benign" course (5). The other siblings, in whom we could detect no V<sub>H</sub>251 rearrangement, required no therapy or exhibited spontaneous remission. If the course of the disease in this family is related to clonal stimulation of V<sub>H</sub>251-bearing cells through either a (neoplastic) T cell or virus that recognizes its associated idiotope, the inability to prevent V<sub>H</sub>251 expression could lead to more severe disease. This predicts that passive immunization with V<sub>H</sub>251 or V<sub>H</sub>WS1 antibody could block stimulation of V<sub>H</sub>WS1-bearing cells and reduce B-cell proliferation.

Few CLL DNAs have been analyzed at the molecular level. Thus we cannot be sure whether the bias seen here is connected to this particular family or to CLL in general. We do know, however, that the V<sub>H</sub>251 gene is not unique to this family, as it has been seen in DNA of over 100 unrelated individuals (data not shown). Now that the V<sub>H</sub>V family of genes has been identified and characterized, the probe can be used on a large number of cases to establish its relationship to the disease in general.

In summary, we have shown that a single V<sub>H</sub> gene belonging to a previously unidentified V<sub>H</sub> gene family was rearranged in two siblings with CLL. More information on the relationship of this V<sub>H</sub> gene family to CLL will depend on determining the distribution of V<sub>H</sub>251 expression in B cells of members of this family and other CLL cases.

We thank Drs. William Blattner and Jeff Clark and Mary Fraser for providing patient cells and cell lines; Drs. Dean Mann and Martin Weigert for helpful discussions; Dr. Charlotte Word for providing patient fibroblast and normal liver DNA; Dr. Oliver Smithies for the leukocyte Charon 35 HindIII library; Dr. Keith Chan and Timothy Durfee for sequencing assistance; Steven Cary and Chkya Das for technical assistance; and Nola Peterson, Marilyn Gardner, and Andrew Hopfensperger for assistance in preparation of this manuscript. This is paper 2909 from the Laboratory of Genetics. This work

was supported by Grants CA 31013 and AI18016 from the National Institutes of Health.

1. Foon, K. & Gale, R. (1985) in *Leukemia: Recent Advances in Biology and Treatment*, eds. Gale, R. P. & Golde, D. W. (Liss, New York), pp. 675-714.
2. Gunz, F. & Veale, A. (1969) *J. Natl. Cancer Inst.* **42**, 517-524.
3. Fraumeni, J., Vogel, C. & DeVita, V. (1969) *Ann. Internal. Med.* **71**, 279-284.
4. Blattner, W., Strober, W., Muchmore, A., Blaese, R., Broder, S. & Fraumeni, J. (1976) *Ann. Internal. Med.* **84**, 554-557.
5. Neuland, C., Blattner, W., Mann, D., Fraser, M., Tsai, S. & Strong, D. (1983) *J. Natl. Cancer Inst.* **71**, 1143-1150.
6. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 280-281.
7. Feinberg, A. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6-13.
8. Rimm, D., Horness, D., Kucera, J. & Blattner, F. (1980) *Gene* **12**, 301-309.
9. Loenen, W. & Blattner, F. (1983) *Gene* **26**, 171-179.
10. Viera, J. & Messing, J. (1982) *Gene* **19**, 771-780.
11. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.
12. Cory, S. & Adams, J. (1980) *Cell* **19**, 37-51.
13. Hieter, P., Maizel, J. & Leder, P. (1982) *J. Biol. Chem.* **257**, 1516-1522.
14. Ravetch, J., Siebenlist, U., Korsmeyer, S., Waldmann, T. & Leder, P. (1981) *Cell* **27**, 583-591.
15. Kabat, E., Wu, T., Bilofsky, H., Reid-Miller, M. & Perry, H. (1983) *Sequences of Proteins of Immunological Interest* (U.S. Dept. of Health and Human Services, Washington, DC).
16. Parslow, T., Blair, D., Murphy, W. & Granner, D. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2650-2654.
17. Mathysens, G. & Rabbits, T. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6561-6565.
18. Rechavi, G., Ram, D., Glazer, L., Zakut, R. & Givol, D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 855-859.
19. Kodaira, M., Kinashi, T., Umemura, I., Matsuda, F., Noma, T., Ono, Y. & Honjo, T. (1986) *J. Mol. Biol.* **190**, 529-541.
20. Kudo, A., Ishihara, T., Mishimura, Y. & Watanabe, T. (1985) *Gene* **33**, 181-189.
21. Takahashi, N., Tetaert, D., Debuire, B., Lin, L.-C. & Putnam, F. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2850-2854.
22. Putnam, F., Florent, G., Paul, C., Shinoda, T. & Shimizu, A. (1973) *Science* **182**, 287-291.
23. Erikson, J., Finan, J., Tsujimoto, Y., Nowell, P. & Croce, C. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4144-4148.
24. Tsujimoto, Y., Jaffe, E., Cossman, J., Gorham, J., Nowell, P. & Croce, C. (1985) *Nature (London)* **315**, 340-343.
25. Yancopoulos, G., Desiderio, S., Paskind, M., Kearney, J., Baltimore, D. & Alt, F. (1984) *Nature (London)* **311**, 727-733.
26. Perlmutter, R., Kearney, J., Chang, S. & Hood, L. (1985) *Science* **227**, 1597-1601.
27. Newkirk, M., Mageed, R., Jefferis, R., Chen, P. & Capra, D. (1987) *J. Exp. Med.* **166**, 550-564.
28. Plater-Zyberk, C., Maini, R., Lam, K., Kennedy, T. & Janossy, G. (1985) *Arthritis Rheum.* **28**, 971-976.
29. Griffiths, G., Berek, C., Kaartinen, M. & Milstein, C. (1984) *Nature (London)* **312**, 271-275.
30. Cleary, M., Meeker, T., Levy, S., Lee, E., Trela, M., Sklar, J. & Levy, R. (1986) *Cell* **44**, 97-106.