A Fetally Expressed Immunoglobulin V_H1 Gene Belongs to a Complex Set of Alleles

Eric H. Sasso,* Ko Willems van Dijk,‡ Andrew P. Bull, and Eric C. B. Milner§
*Virginia Mason Research Center, Seattle, Washington 98101; and *Division of Rheumatology and Departments of ‡Pathology and §Immunology, University of Washington, Seattle, Washington 98195*

Abstract

The immunoglobulin V_H gene 51p1, a member of the large V_H1 gene family, is preferentially expressed by B cells in the fetus and in chronic lymphocytic leukemia (CLL) and appears to be the source for many cryoglobulin rheumatoid factors. Polymorphism of 51p1 may therefore be functionally important. We have studied the germline representation of 51p1 and closely related V_H elements to establish their prevalence and allelic relationship. A panel of oligonucleotide probes directed to the complementarity determining regions (CDR1 and CDR2) of 51p1 and a similar gene, hv1263, was used in restriction fragment polymorphism analysis of 48 unrelated individuals and six families. 13 V_H alleles to the 51p1 locus were identified, each distinguished by its restriction fragment size, hybridization profile, or both. On some haplotypes the locus was duplicated. Null alleles were not seen. The 13 alleles were cloned, yielding nine distinct nucleotide sequences that were > 98.2% identical and included 51p1 and hv1263. These germline variations could influence specificity for antigen, because the corresponding protein sequences differed by up to five amino acids, including three nonconservative changes in the CDR. Two of the most prevalent variants contained 51p1. These findings expand the spectrum of polymorphism seen among human V_H genes and elucidate the germline origin of V_H1 sequences frequently expressed in autoantibodies and CLL. We conclude that the 51p1 locus is polymorphic, and that the 51p1 element is the predominant member of a complex set of alleles. (*J. Clin. Invest.* 1993. 91:2358–2367.) Key words: antibody diversity • genes, immunoglobulin V_H germline • molecular sequence data • oligonucleotide probes • restriction fragment length polymorphism

Introduction

The variable regions of human Ig heavy chains are encoded by somatically rearranged segments selected from the V_H, D, and J_H gene clusters on chromosome 14. Human V_H elements can be divided into at least seven gene families, where members of the same V_H family share at least 80% nucleotide sequence identity (1–9). Gene segments from different human V_H families are interspersed throughout the V_H cluster, which contains upwards of 100 V_H loci per haploid genome (3, 6). Recent studies indicate that many human V_H genes are polymorphic, presumably the result of extensive duplication, deletion, gene conversion, and point mutation in the V_H cluster (10–21). The repeated cloning of some V_H sequences from unrelated individuals (4–8, 22–27) and the results of sequence-specific hybridization studies (19–21) suggest that polymorphic V_H loci are commonly dominated by only a few conserved alleles (i.e., two to four) (20). However, because linkage disequilibrium has been found among only a minority of the studied V_H elements and thus appears to be limited to elements of physically proximal V_H loci the diversity of V_H haplotypes in the population is large (20, 28).

Recent studies indicate that the physical order and location of loci in the human V_H cluster is similar among haplotypes, regardless of the particular polymorphisms the haplotypes contain (27, 29, 30). Allelic relationships among V_H elements should be discernible, therefore, if one can account for the deletions and duplications affecting the corresponding locus. Hybridization studies have identified paired restriction fragment length polymorphisms of V_H elements that appear to be related as alleles (21, 28). A more complete understanding of allelism among V_H genes has been slow to emerge, however, because the complex polymorphism of V_H genes and the insensitivity of Southern hybridization to small sequence variations have hindered precise correlation of Southern hybridization bands with known V_H sequences. Thus, the identity of V_H elements contained in allelic RFLP is not usually known, and, when similar V_H sequences are cloned from unrelated individuals, it is not routinely possible to determine whether they are alleles or derived from distinct loci.

We have previously identified V_H polymorphism by performing genomic hybridization with oligonucleotide probes (19–21). This approach confers exquisite sequence specificity upon RFLP analysis and has revealed a spectrum of polymorphism among genes from the V_H3 and V_H4 families (19–21, 31). In the present study, we examine the polymorphism of 51p1 and hv1263, similar genes from the large V_H1 family. The 51p1 gene is expressed in the restricted repertoire of fetal liver B lymphocytes, and as such might influence the development of the adult V_H repertoire (8). Serological evidence implicates 51p1 or related V_H1 elements as the rearranged V_H gene in about one-quarter of chronic lymphocytic leukemia (CLL)[1] B cell lines (32) and as a candidate germline source for many paraprotein rheumatoid factors (33). It is not known if all individuals possess the 51p1 germline gene, if there exist alleles of 51p1, or if such alleles are equal participants in the fetal repertoire, B cell neoplasia, or autoimmunity. It is thus important to understand the germline representation of 51p1 because poly-

Address correspondence and reprint requests to Eric C. B. Milner, Virginia Mason Research Center, 1000 Seneca St., Seattle, WA 98101.

Received for publication 24 August 1992 and in revised form 28 December 1992.

J. Clin. Invest.

0021-9738/93/06/2358/10 $2.00
Volume 91, June 1993, 2358–2367

1. *Abbreviations used in this paper:* CLL, chronic lymphocytic leukemia; CDR, complementarity determining region.

understand the germline representation of 51p1 because polymorphism at the 51p1 locus might influence the expressed B cell repertoire in normal and abnormal states. We now present results of allele-specific hybridization analysis and gene cloning. It is demonstrated that 51p1 belongs to a complex set of V_H alleles, the definition of which establishes the germline origin of many V_H1 sequences expressed in the fetus, in CLL, and in autoantibodies.

Methods

Enzymatic digestion of genomic DNA. DNA was prepared by standard methods from peripheral blood leukocytes obtained from 48 unrelated healthy donors and from members of two families (Virginia Mason Research Center families 117 and 292) (34). One member from one family (subject 3116, offspring in family 117) was included in the population study. DNA was similarly prepared from EBV-transformed B cells of members of four additional families (EYo, ECa, GAW0.4, GAW0.3), two of which (EYo and ECa) have been previously described (19). Digestions were performed with 3–4 U TaqI/μg genomic DNA followed by extraction with phenol and chloroform/isoamyl alcohol, ethanol precipitation, and resuspension in TAE (0.04 M Tris-acetate, 0.002 M EDTA) running buffer.

Oligonucleotide probes. Synthetic deoxyoligonucleotide probes were prepared at the Howard Hughes Medical Institute at the University of Washington, Seattle, WA. Probe M28 (codons 54–60, numbering system of Kabat et al. (35), in complementarity determining region 2 [CDR2]) matches a sequence from the V_H germline gene hv1263 (33), 5′ CTT GGT ATA GCA AAC TAC GCA 3′. Probes matching sequence from the cDNA V_H element 51p1 (8) are M27 (codons 54–60 in CDR2), 5′ TTT GGT ACA GCA AAC TAC GCA 3′; H111 (codons 50–56 in CDR2) 5′ GGG ATC ATC CCT ATC TTT GGT 3′; and M32 (codons 70–76 in FR3) 5′ ACC GCG GAC GAA TCC ACG AGC 3′. Probe M26 (codons 30–36 in CDR1) 5′ AGC AGC TAT GCT ATC AGC TGG matches a sequence that is identical in hv1263 and 51p1. Antisense probes M116, M117, and M118 are the reverse complements of H111, M27, and M28, respectively.

Hybridization with oligonucleotide probes. Hybridization of ^{32}P-labeled oligonucleotide probes to DNA in dried agarose gels was performed according to a previously described protocol (20). Hybridization temperatures were 55–57°C for probes M26, M27, M28, and H111 and 52°C for probe M32. After hybridization, gels were washed for 1 h at 43°C in 5× SSC, 0.5% SDS, and then postwashed for 1 h in 3.2 M tetramethylammonium chloride, 0.5% SDS, 0.001 M EDTA, and 0.05 M Tris, pH 7.5. Temperatures of postwashes are given in the figure legends. The predicted melting temperature (T_m) of hybridized oligonucleotide probes in the tetramethylammonium chloride postwash solution, determined by probe length, is approximately 62°C for all oligonucleotide probes used in this study (36). Hybridization to colony lifts on nitrocellulose filters followed the same protocol except that prehybridization was performed with hybridization buffer, without probe, and this buffer contained 100 μg/ml boiled salmon sperm DNA, in addition to the previously described components. Antisense probes were used only for filter hybridization at 55–57°C. Filters were postwashed at 55–59°C.

Film exposure of hybridized gels. After postwashing, gels were blotted dry on paper towels, placed on filter paper backing, wrapped in cellophane, and exposed to Kodak X-OMAT AR film (Eastman Kodak Co., Rochester, NY) in a sealed metal cassette with an intensifying screen at −70°C for the durations indicated in the figure legends.

Reuse of gels. After hybridization and film exposure, gels were prepared for the next hybridization by incubation for 1 h in 0.5 M NaOH, 1.5 M NaCl, followed by two successive 30-min washes in 3.0 M NaCl, 0.5 M Tris-HCl, pH 7.0, and a brief wash in 0.01 M Tris, pH 8.0, all at room temperature.

Cloning of selected V_H germline genes. 1-μg aliquots of genomic DNA prepared from peripheral blood leukocytes of nine subjects were amplified by 30 cycles of PCR using the Geneamp kit (Perkin-Elmer Cetus Instruments, Norwalk, CT). Each cycle was denatured for 1 min at 93°C, annealed for 2 min at 55°C, and extended with TaqI polymerase for 3 min at 72°C, with a 7-min extension after the last cycle. PCR primers were TTC TTG GTG GCA GCA GCC ACA GG (5′ primer) and AG GAT GTG GTT TCT CAC ACT GTG (3′ primer), corresponding to the hv1263 sequence (33) immediately 5′ of the leader intron and 3′ of the V_H coding region, respectively. PCR product (1–4 μl; in some cases, first separated in an agarose gel, extracted, and ethanol precipitated) was blunt-end ligated into the HincII or SmaI site of M13mp19 (Pharmacia LKB Biotechnology Inc., Uppsala, Sweden), transformed into DH5αF′ competent cells, and plated in YT agarose with JM101 lawn cells, isopropyl-β-D-thiogalactopyranoside, and 5-bromo-4-chloro-3-indoyl-β-D-galactoside according to standard methods. Nitrocellulose filter lifts of the primary plates were screened by hybridization with appropriate sense and antisense oligonucleotide probes (Table I) to identify clones with inserts in each orientation. All desired primary clones were picked, diluted, and replated, and filter lifts of these secondary plates were screened with the appropriate oligonucleotide probes.

Nucleotide sequence analysis of selected V_H germline genes. Sequence analysis (Sequenase 2 reagents and protocol; U.S. Biochemical Corp., Cleveland, OH) was performed only with DNA prepared from selected M13 secondary clones to exclude undesired V_H genes that could have been inadvertently picked alongside a primary clone of interest. Clones in each direction were sequenced for all gene variants, except those from subject 3027. For 44 of the 45 primary clones sequenced, all 392 nucleotides between the PCR primers were determined (1 clone had 326 nucleotides). Each of the 15 sequences listed in Table II reflects a consensus of two to five primary clones whose nucleotide sequences were identical, except for unique substitutions appearing on only a single clone, which were presumed to be PCR errors and excluded from the reported consensus sequence. Nine such substitutions were excluded. For the two sequences obtained from subject 3116, the same protocol was followed except that: (*a*) template DNA for PCR amplification was prepared from two lambda phage clones selected by screening a previously described library of subject 3116 genomic DNA (31) with probe M27, (*b*) antisense M13 clones were identified by screening with a V_H1 cloned probe (51p1) to detect M13 plaques not seen with the sense probe M27, and (*c*) DNA for sequencing reactions was prepared from M13 primary clones.

Results

To identify the V_H element 51p1 and its potential alleles, genomic hybridization was performed with four oligonucleotide probes specific for the CDR of 51p1 and the V_H gene hv1263, whose coding regions differ by only five nucleotides. Two of the CDR2 probes, M27 and M28, detected mutually exclusive sets of V_H elements because of the nucleotide mismatches distinguishing their respective targets in 51p1 and hv1263 (Fig. 1). These two sets of V_H elements are designated 51p1 related ($M27^+$) and hv1263 related ($M28^+$).

Detection of RFLP containing hv1263-related germline genes. When TaqI-digested genomic DNA from 48 unrelated individuals was hybridized with the hv1263-specific probe M28, 58% of individuals were positive and 42% were blank. The representative gel in Fig. 2 *A* shows examples of each of the four observed $M28^+$ restriction fragments, at 5.8 kb (lanes *4* and *7*), 4.3 kb (lane *11*), 4.0 kb (lanes *5* and *8*), and 3.9 kb (lanes *1, 3, 4, 6, 9, 10, 12,* and *13*).

In two individuals, probe M28 detected two bands (e.g., lane *4*) and, in five others, an intense single band, consistent with a double gene dose (e.g., lanes *1, 3,* and *8*). In no case did M28 detect three bands or an intense band plus a second band,

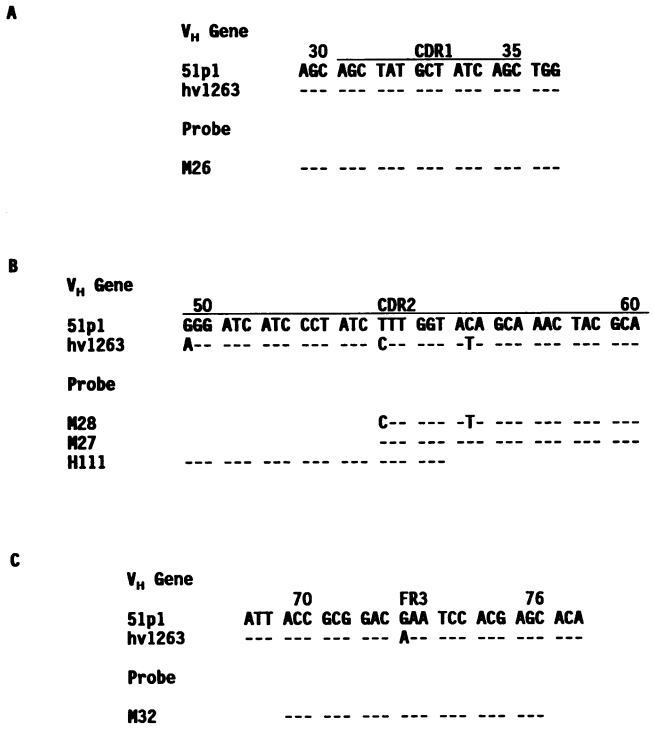

Figure 1. Sequence-specific oligonucleotide probes. Nucleotide sequences of V_H1 target genes 51p1 and hv1263 are depicted in regions of (*A*) CDR1, (*B*) CDR2, and (*C*) FR3, to which oligonucleotide probes M26 (*A*), M28, M27, H111 (*B*), and M32 (*C*) are directed. Nucleotides of hv1263 and of all oligonucleotide probes are specified only where they differ from gene 51p1. Residues comprising CDR1 and part of CDR2 are indicated by a horizontal line over 51p1. Codon numbering is according to Kabat et al. (35).

suggesting that two M28$^+$ elements is the usual maximum occurring in a genome.

Detection of RFLP containing 51p1-related germline genes. When TaqI-digested genomic DNA from the same 48 unrelated individuals was hybridized with the 51p1-specific probe M27, 85% of individuals were positive and 15% were blank. The representative gel in Fig. 2 *B* shows examples of each of the three most prevalent restriction fragments, at 7.5 kb (lanes *5, 7, 9, 10,* and *14*), 4.3 kb (lanes *2, 6* and *11–14*), and 4.0 kb (lanes *2, 5, 7, 9, 10,* and *14*). In addition, infrequently occurring bands were detected by probe M27 at 6.5 and 3.9 kb (not shown).

In four individuals, M27$^+$ bands were increased in intensity, suggesting a double gene dose. Two of these individuals possessed intense 7.5- and 4.0-kb bands; two others possessed an intense 4.3-kb band (no examples shown). In this population, no individual with an intense band detected by either M28 or M27 had any bands detected by the other of these two probes.

Detection of CDR2 sequence variants among 51p1-related germline genes. To determine whether nucleotide sequences of 51p1-related elements vary in the region of CDR2 located 5′ to that targeted by probe M27, genomic hybridization of the same 48 individuals was performed with a second 51p1-specific probe, H111 (Fig. 1). With only two exceptions, all bands detected by probe M27 were also detected by probe H111 (Fig. 2 *C*). In the two exceptions, an M27$^+$ band was not detected by probe H111, once at 4.3 kb (not shown) and once at 4.0 kb (Fig. 2, *B* and *C*, lane *2*). These two M27$^+$H111$^-$ elements must therefore be sequence variants that differ from the 51p1 sequence at one or more of the nucleotide positions uniquely targeted by probe H111. In no case did probe H111 detect an element not also detected by probe M27. This finding, that probes M27 and H111 detect identical patterns of RFLP in most individuals, indicates that the 51p1-related elements we have detected possess little sequence variation in the CDR2 regions targeted by probes M27 and H111.

Detection of CDR1 sequence variants among 51p1-related germline genes. To determine whether nucleotide sequences of 51p1-related genes vary in CDR1, genomic hybridization of the same 48 individuals was performed with the CDR1 probe M26 (Fig. 1). With one exception, all 51p1-related elements previously detected by CDR2 probe M27 were also detected by probe M26. The exception was the single M27$^+$H111$^-$ band at 4.0 kb described above (Fig. 2 *B–D,* lane *2*). As expected, probe M26 also detected hv1263-related elements seen with the CDR2 probe M28$^+$ (see below). Thus, in the present study population, nearly all 51p1-related elements appear to have conserved 51p1 sequences in CDR1 as well as CDR2.

Detection of CDR1 sequence variants among hv1263-related germline genes. Hybridization with the CDR1 probe M26 detected most of the hv1263-related elements seen previously with the CDR2 probe M28, but 5 of 30 M28$^+$ bands were blank with probe M26 (e.g., Fig. 2, *A* and *D,* lanes *6* and *8*). In addition, some M26$^+$ bands were less intense than expected if M26 were to detect both the M28$^+$ and the M27$^+$ restriction fragments known to comigrate or migrate similarly in those individuals (e.g., Fig. 2 *D,* compare the 4.0/3.9 kb band in lane *9* with the 4.0 kb band in lane *10*). In such cases, the implication is that M26 detected an M27$^+$ element, but not the M28$^+$ element. These results indicate that although many of the hv1263-related elements have conserved CDR1 as well as CDR2, several CDR1 variants exist. The hybridization results

Table I. Related V_H1 RFLP

Variant	kb	Probes M26	Probes H111	Probes M27	Probes M28	Prevalence
1	7.5	+	+	+	−	0.50
2	6.5	+	+	+	−	0.02
3	5.8	+	−	−	+	0.08
4	4.3	+	−	−	+	0.02
5	4.3	+	+	+	−	0.58
6	4.3	+	−	+	−	0.02
7	4.0	+	+	+	−	0.48
8	4.0	−	−	+	−	0.02
9	4.0	+	−	−	+	0.04
10	4.0	−	−	−	+	0.17
11	3.9	+	+	+	−	0.06
12	3.9	+	−	−	+	0.23
13	3.9	−	−	−	+	0.13

The TaqI restriction fragment sizes (kb), oligonucleotide probe hybridization profiles, and phenotypic prevalences of 13 distinct hybridization bands are listed in order of their assigned variant numbers. Each variant contains a 51p1-related (M27$^+$) or hv1263-related (M28$^+$) gene. Data were obtained by sequential hybridization of the probes to genomic DNA from 48 subjects. When a subject possessed a variant with increased band intensity, the variant was counted once.

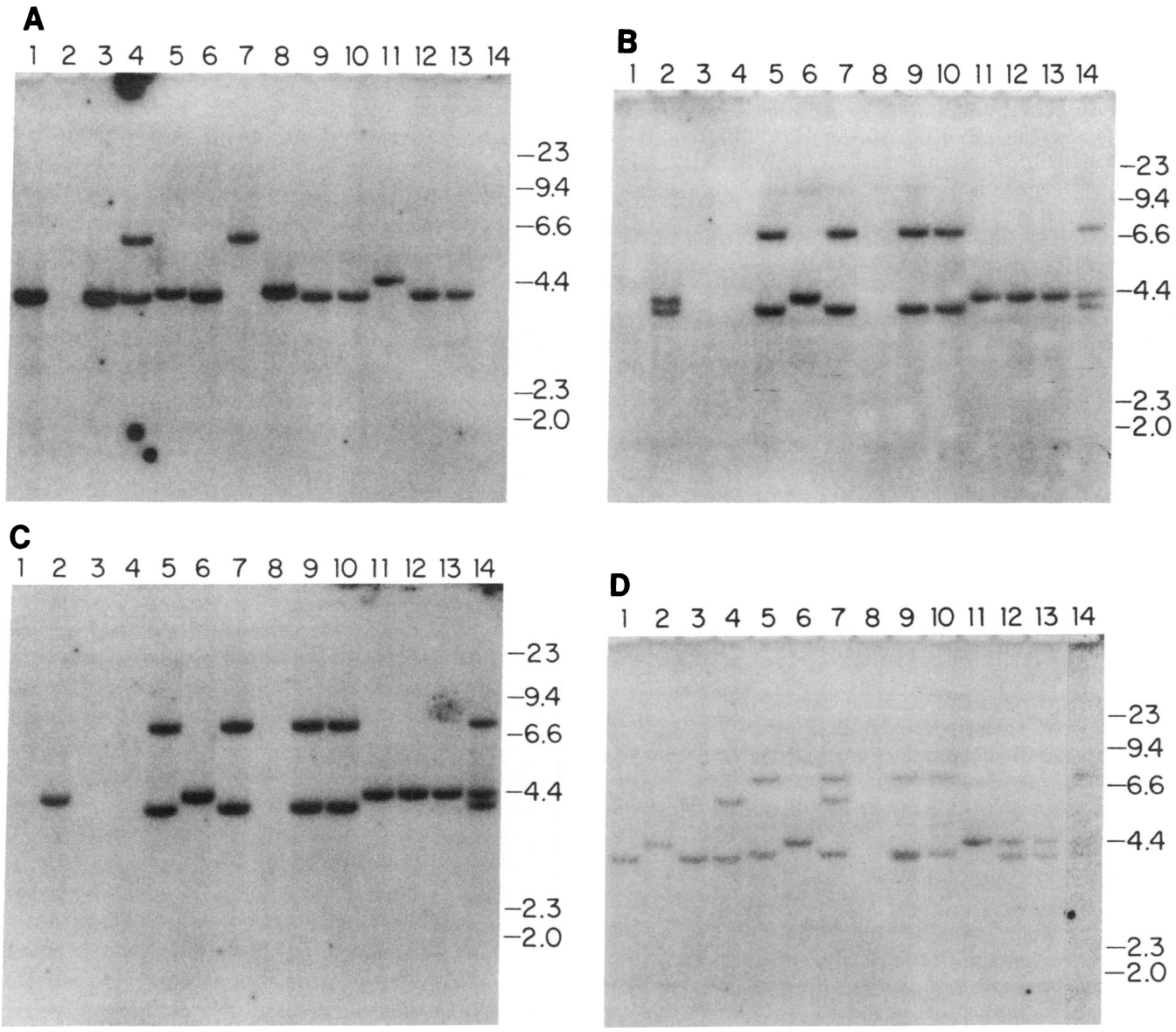

Figure 2. Identification of 51p1 and related germline genes. DNA extracted from PBL of 14 unrelated individuals was digested with TaqI and separated on two identical 1% agarose gels, which were then dried and prepared for hybridization. One gel was sequentially hybridized with oligonucleotide probes M28 (*A*), M27 (*B*), and M26 (*D*). The other gel (*C*) was hybridized with probe H111. Postwash temperatures were 57°C in *A, B,* and *D,* and 58°C in *C.* All films were exposed for 7 d.

do not establish whether the CDR1 sequences of the variants are identical to each other.

In no case did probe M26 detect elements previously undetected by either M28 or M27 (Table I). Thus, the CDR1 sequence shared by hv1263 and 51p1 is not found on any germline genes other than the hv1263- and 51p1-related elements described above.

13 distinct V_H elements are related to hv1263 and 51p1. The results of genomic hybridization with probes M28, M27, H111, and M26 identified 13 distinct V_H germline elements on the basis of differences in restriction fragment size, oligonucleotide hybridization profile, or both (Table I). Elements detected by all three 51p1-derived probes (i.e., $M26^+$, $H111^+$, $M27^+$) occurred on five different restriction fragments, of which three (7.5, 4.3, and 4.0 kb) were prevalent, occurring in 48–58% of individuals (Table I). Elements detected by both hv1263-derived probes (i.e., $M26^+$, $M28^+$) occurred on four restriction fragments, of which one, at 3.9 kb, was relatively prevalent, occurring in 23% of individuals (Table I). Also identified were four CDR sequence variants, with prevalences of 2–17% (Table I).

Apparent duplication of a 51p1-related gene. The ($M26^+$, $M27^+$, $H111^+$) bands at 7.5 and 4.0 kb, variants 1 and 7, respectively (Table I), cooccurred in 22 individuals (e.g., Fig. 2, lanes *5, 7, 9, 10,* and *14*) and occurred separately in only three individuals (not shown). By random assortment, cooccurrence of these 7.5- and 4.0-kb bands was expected in 12 individuals. This result indicates linkage disequilibrium between these two elements ($\chi^2 = 9.59$, $P < 0.0001$) and suggests they represent a recent duplication.

Chromosomal organization of genes related to 51p1 and hv1263. To examine the chromosomal organization of the V_H

elements related to 51p1 and hv1263 and to determine whether the two elements of 51p1-related 7.5/4.0-kb pair segregate together, as their distribution in the population predicts, six two-generation families were studied by genomic hybridization with probes M28 (hv1263), M27 (51p1), and H111 (51p1). In four families (292, EYo, GAW0.4, and GAW0.3), only one parent had the 7.5/4.0-kb pair detected with M27 or H111. In all four of these families, each offspring inherited from that parent either both bands of the 7.5/4.0-kb pair or the other band possessed by that parent (Fig. 3). In a fifth family, EHa, both parents possessed the 7.5/4.0-kb pair detected with H111. The other maternal element in the EHa family, an HIII$^+$ 4.3-kb band, appeared in all offspring, suggesting that each offspring received its 7.5/4.0-kb pair from the homozygous father (Fig. 3). These data indicate that the two elements of the 7.5/4.0-kb pair, when present, reside on the same chromosome as each other and on the opposite chromosome of any other element detected in the same genome by the CDR2 probes M28, M27, and H111.

In no family was a null haplotype observed (Fig. 3). In five of the studied families, parental chromosomes detected by M28, M27, or H111 contained either a single band or the 7.5/4.0-kb pair. In the sixth family, 292, one of the maternal chromosomes contained two distinct elements different from the 7.5/4.0-kb pair (i.e., an M27$^+$ 6.5-kb band plus an M28$^+$ 4.0-kb band) (Fig. 3). This gene pair likely represents a rare linkage of two V_H1 elements that usually occur on opposite chromosomes. Thus, the segregation data demonstrate that, in the families studied, V_H haplotypes contain either one of the elements detected by the CDR2 probes or a linked pair of these elements, usually the 7.5/4.0-kb 51p1-related pair.

Allelic distribution of genes related to hv1263 and 51p1. When the bands detected by probes M28 and M27 are analyzed in each of the 48 individuals studied, 46 individuals could be interpreted to be: (*a*) heterozygous, with two different bands, one per haplotype (e.g., Fig. 2, *A* and *B*, lanes *2, 4, 6, 11–13*), (*b*) heterozygous, with the M27$^+$ 7.5/4.0-kb pair on one haplotype and a single distinct band on the other (e.g., Fig. 2, *A* and *B*, lanes *5, 7, 9, 10,* and *14*), or (*c*) homozygous, with the same hybridization band (or the M27$^+$ 7.5/4.0-kb pair) on each haplotype (e.g., Fig. 2, *A* and *B*, lanes *1, 3,* and *8*). The two remaining individuals (of 48) possessed three distinct elements, but not the 7.5/4.0-kb pair (subject 3050, variants 1, 5, and 9; subject 3110, variants 2, 6, and 12; both not shown), indicating that, other than variants 1 and 7 in the M27$^+$ 7.5/4.0-kb pair, duplications among these related V_H elements are not common in the population studied (i.e., 2 of 96 haplotypes).

The distribution of these elements in the population and in the six families is entirely consistent with that expected of a set of alleles to a single locus, if the apparent duplications, such as the 7.5/4.0-kb pair, are regarded as one allele plus a closely linked V_H element. In the population of 48 individuals, no genome contained the 7.5/4.0-kb pair (variants 1 and 7, prevalence = 0.46) with the (M26$^+$, M27$^+$, H111$^+$) 4.3-kb band (variant 5, prevalence = 0.58) and any of the M28$^+$ elements (variants 3, 4, 9, 10, 12, and 13, prevalence = 0.67). By random assortment, such hybridization patterns were expected in at least eight individuals ($\chi^2 = 8.51$, $P < 0.01$).

Sequence analysis of V_H variants related to hv1263 and 51p1. The nucleotide sequences were determined for each of the 13 variants identified by genomic hybridization (Table II).

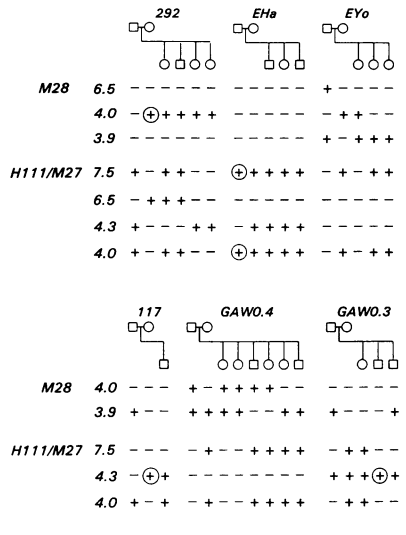

Figure 3. Segregation in families of 51p1- and hv1263-related germline elements. Genomic DNA from PBL or EBV-transformed B cell lines from subjects in six two-generation families were digested with TaqI and separated on 1% agarose gels, which were dried down, denatured, and neutralized. Gels were sequentially hybridized with probes to hv1263 (M28) and 51p1 (M27 or H111). Postwashes were performed at 53–57°C. The detected V_H elements are indicated by a +, according to the hybridizing probe and the restriction fragment size (kb). Hybridization bands of increased intensity are indicated by a + contained in a circle. Family code names appear above each family diagram.

A total of nine different nucleotide sequences was obtained (Fig. 4). The sequences have identical leader introns, with one exception, and differ at only eight coding region positions, making them 98.2–100% identical to each other. Four sequences, 027M27A, 079T8M27, 113T4M32, and 1d144 were

Table II. Sequencing of V_H1 Hybridization Variants from Selected Genomes

Subject	Variant	Sequence
3027	5	027M27A (51p1)
	11	027M27B
3079	1	079T8M27 (51p1)
	9	079T4M28 (hv1263)
3110	2	110T6M27
	6	110T4M27
	12	110T4M28
3113	7	113T4M27
	1/5	113T4M32 (51p1)
3116	5	1d144 (51p1)
	8	1d194
3121	3	121T5M28 (hv1263)
	1	Not sequenced
	7	Not sequenced
3122	10	122T4M28
3124	13	124T4M28
	1	Not sequenced
	7	Not sequenced
3125	4	125T4M28
	5	Not sequenced

Germline V_H1 sequences were cloned from the indicated subjects. For each subject, all variants (as defined in Table I) are indicated, as are the names of the corresponding nucleotide sequences cloned from that subject. Sequence 113T4M32 could have been derived from variant 1 and/or variant 5, both of which were present in subject 3113. Not sequenced signifies that no attempt was made to sequence clones containing that variant from that individual. Nucleotide sequences identical to those of 51p1 or hv1263 are indicated.

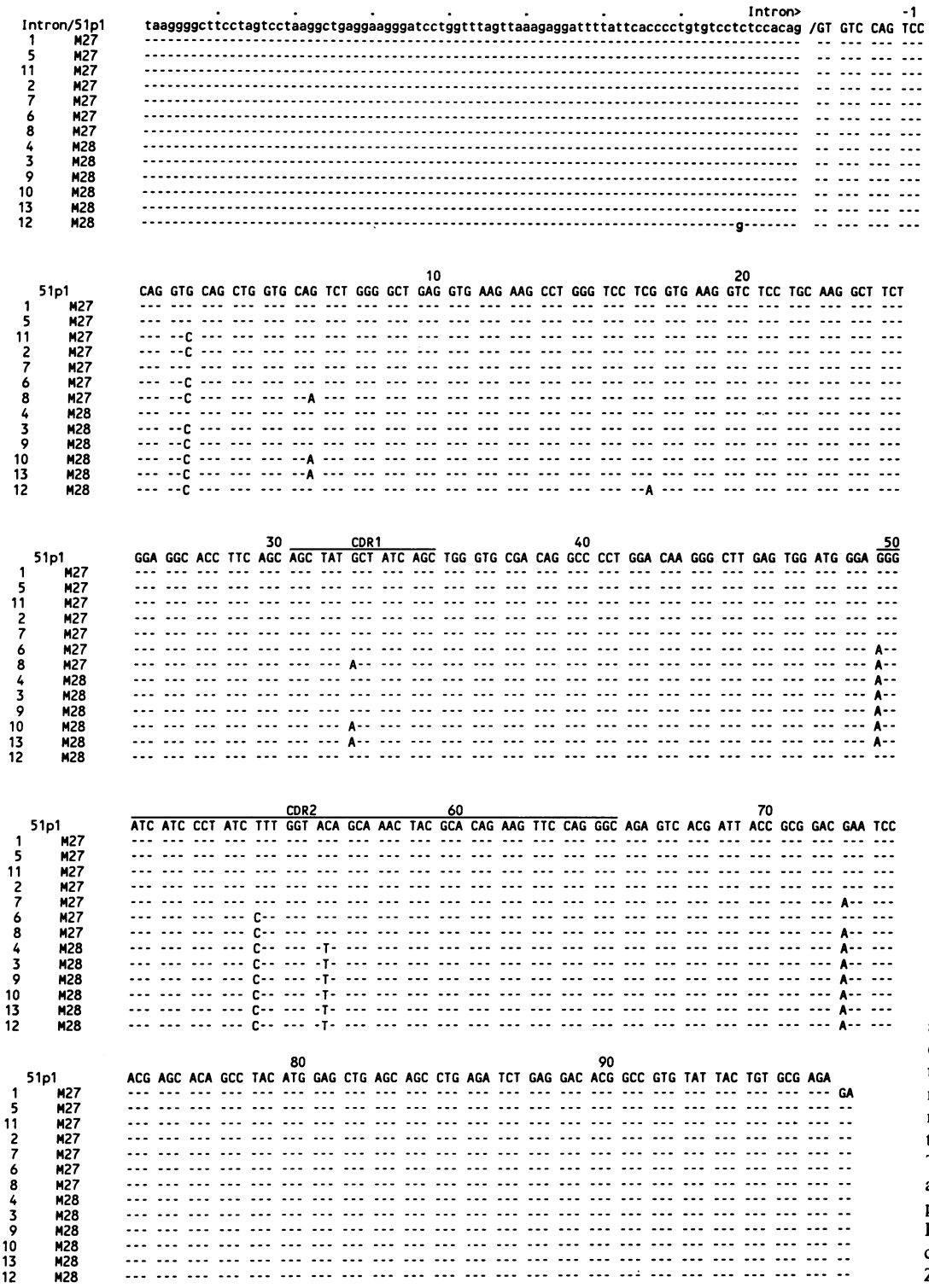

Figure 4. Sequence analysis of the 51p1 allelic set. Nucleotide sequences of the 13 V_H1 hybridization variants summarized in Table I were determined by selectively cloning the genes from genomic DNA of nine different subjects (Table II). The corresponding hybridization variant numbers (1–13) and the detecting CDR2 probes (M27 or M28) appear next to each sequence. The 51p1 germline sequence (variants 1 and 5), which is identical in its coding region to the available sequence from cDNA 51p1 (8) and in its leader intron to that of germline gene hv1263 (33) appears above as a reference sequence. Other sequences are specified only where different from the reference. Codon numbering and CDR demarcation are according to Kabat et al. (35). These sequence data are available from the European Molecular Biology Laboratory under the accession numbers Z14295–Z14309.

identical to 51p1. The nine nucleotide sequences encode seven distinct polypeptides (Fig. 5). Amino acid replacements result from five of the nucleotide substitutions, of which four are in the CDR. Nonconservative amino acid substitutions occur at codons 33 in CDR1, 50 and 56 in CDR2, and 73 in FR3.

Of 17,574 nucleotides sequenced, 9 suspected PCR artifacts were encountered, none of which corresponded to the substitutions described above. All were unique and all but two were outside the CDR. Four of the nine suspected artifacts were T to C or A to G changes, the most common type of Taq polymerase error (37, 38).

Assignment of nucleotide sequences to genomic restriction fragments. The assignment of obtained sequences to the correct restriction fragment variants in Table I was in most cases straightforward because the DNA samples selected for cloning were from subjects whose 51p1- or hv1263-related variants were identifiable by distinct combinations of oligonucleotide probes. For example, in subject 3116, genomic hybridiza-

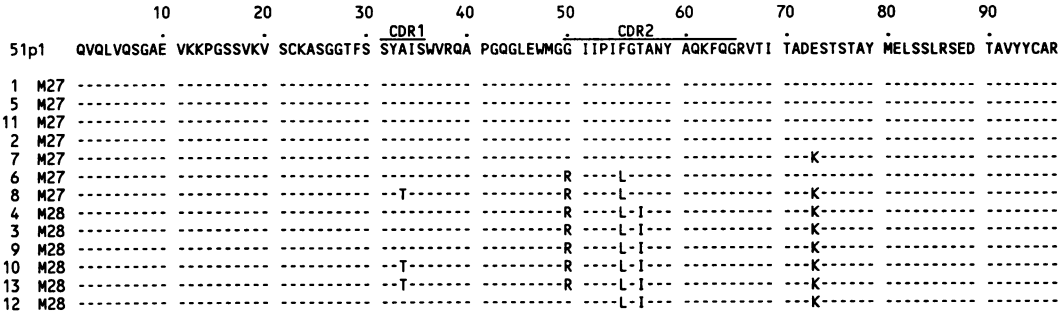

Figure 5. Amino acid sequences of the 51p1 allelic set. The deduced amino acid sequences of the V_H1 elements appearing in Fig. 4 are displayed beneath the translated amino acid sequence of 51p1. Amino acids are specified only at residues that differ from the 51p1 protein. The corresponding DNA hybridization variant numbers (1–13) and the detecting CDR2 probe (M27 or M28) appear next to each sequence. Codon numbering and CDR demarcation are according to Kabat et al. (35).

tion detected variant 5 ($M27^+$, $H111^+$) and variant 8 ($M27^+$, $H111^-$) (Fig. 2, lane *2*). The two 51p1-related sequences obtained from subject 3116, 1d144 and 1d194 (Table II), were derived from clones having the hybridization profiles ($M27^+$, $H111^+$) and ($M27^+$, $H111^-$), respectively. Thus, variant 5 must contain sequence 1d144 and variant 8 must contain sequence 1d194. The other DNA used for cloning that appear in Fig. 2, from subjects 3121 (lane *7*), 3122 (lane *8*), and 3125 (lane *11*), are similarly illustrative (Fig. 4 and Tables I and II).

Only variant 7 was less straightforward, because this $M27^+$ 4.0-kb band only appeared in genomes containing other V_H elements with the same hybridization profile, ($M26^+$, $H111^+$, $M27^+$). With the DNA chosen for the cloning of variant 7, from subject 3113, genomic hybridization previously showed ($M26^+$, $H111^+$, $M27^+$) bands at 7.5, 4.3, and 4.0 kb (i.e., variants 1, 5, and 7). Only two nucleotide sequences were cloned from subject 3113, however. One sequence, designated 113T4M32, derived from five M13 clones whose sequences were identical to that of 51p1 (Table II). The other, designated 113T4M27, derived from two M13 clones whose sequences each differed from 51p1 by only a single FR3 substitution (Table II, Fig. 4). The 51p1 sequence of subject 3113 could have derived from variants 1 or 5, or both, as was found in subjects who possess variant 1 or 5, but not both (3027, 3079, 3116; Table II). It is deduced, therefore, that the non-51p1 sequence from subject 3113 must have derived from variant 7. Confirmation that the genes on the 7.5- and 4.3-kb fragments (variants 1 and 5) contain the same FR3 sequence and that the gene on the 4.0-kb fragment (variant 7) does not was obtained by genomic hybridization with a 51p1-specific FR3 probe, M32 (Fig. 1), to TaqI-digested DNA of subject 3113 and the 14 individuals shown in Fig. 2. Probe M32 detected all bands that were detected by probe M27 at 7.5 and 4.3 kb but none at 4.0 kb (data not shown). The 7.5/4.0-kb $M27^+$ pair therefore consists of an exact 51p1 sequence at 7.5 kb (variant 1) and a linked FR3 variant of 51p1 at 4.0 kb (variant 7). These assignments of sequences to hybridization bands indicate that in the population of 48 people, 83% possessed restriction fragments shown to contain the 51p1 nucleotide sequence, as variant 1, variant 5, or both, and 96% possessed at least one of the other 11 variants.

Genomic hybridization accurately discriminates nucleotide sequence variations. Each of the obtained nucleotide sequences corroborates the results of genomic hybridization to the donor DNA. These correlations are not an artifact of clone selection because most clones were identified by screening with only one CDR2 probe. All hybridization blanks were explainable by nucleotide mismatches in the cloned target sites. The three alleles not detected by the CDR1 probe M26, variants 8, 10, and 13 (Table I), contained the same G to A substitution in codon 33 (Fig. 4).

In nearly all instances, the hybridization target sites in a cloned gene and the probes that detected the gene in genomic DNA had identical nucleotide sequences. The only two exceptions occurred in sequences 110T4M27 and 1d194 assigned, respectively, to the $M27^+$ variants 6 and 8 (Table II). These sequences each mismatched the 5′ terminal nucleotide of the detecting probe, M27 (compare Figs. 1 and 4). As previously demonstrated for similar results with other V_H genes (20, 31), this finding is consistent with the genomic hybridization data and the results of clone screening because the postwashing steps of each were performed at a submaximal stringency that should have tolerated a terminal nucleotide mismatch. Thus, when hybridization with probe M27 was postwashed at maximal stringency, requiring a perfect match between probe M27 and target DNA, the 4.0-kb restriction fragment containing the sequence of 1d194 in variant 8 was no longer detected (Fig. 6).

Discussion

In this study, we have identified a set of closely related V_H germline elements that includes the fetally expressed gene 51p1. Our approach was to use sequence-specific genomic hybridization to detect genes containing the CDR sequences of 51p1 or a nearly identical gene, hv1263. A total of 13 hybridization variants was found. The chromosomal organization and distribution in the population of these variants is consistent with the conclusion that they comprise an essentially complete set of alleles to the 51p1 gene. The sequences of the 51p1 allelic set differ by a limited number of nonconservative substitutions that could have functional significance. These findings indicate that the 51p1 gene is polymorphic. The implication of this allelic variation is that individual differences at the 51p1 locus may be biologically important.

The 13 hybridization variants we identified reside on six different TaqI restriction fragments and, on the basis of distinct hybridization profiles with a panel of sequence-specific probes, include several different coding region sequences. Examples were seen of RFLP and of coding region sequence variations that existed independently and that occurred together. The three most prevalent variants, 1, 5, and 7 (Table I) had sequences identical or 99.7% identical to 51p1. They each oc-

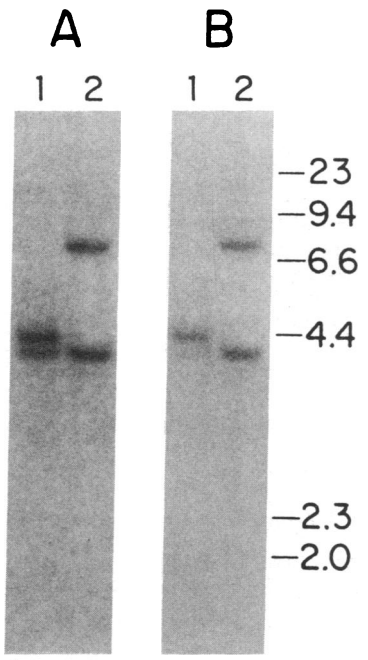

Figure 6. Confirmation that a nucleotide variation revealed by sequence analysis exists in germline DNA. Genomic DNA from PBL of two unrelated individuals (lanes *1* and *2*, subjects 3116 and 3124) were digested with TaqI and separated on a 1% agarose gel. Both subjects are described in Table II. Subject 3116 is the source of sequence 1d194, which differs from probe M27 by a single terminal nucleotide. The gel was dried and hybridized with probe M27. (*A*) After hybridization, the gel was postwashed at 57°C (submaximal stringency) and exposed to film for 4 d. (*B*) The same gel as in *A*, to which probe M27 was still hybridized, was postwashed a second time at 61.5°C (maximal stringency) and exposed to film for 8 d. The band at 4.0 kb in lane *1* has preferentially washed off, consistent with a terminal mismatch between the target sequence it contains, 1d194, and probe M27.

curred in about half the population and resided on different TaqI fragments. The next most prevalent was variant 12 (Table I), which occurred in about one-quarter of individuals and contained a unique nucleotide sequence encoding a protein similar to that of hv1263. The exact hv1263 nucleotide sequence occurred on two restriction fragments present in 8 and 4% of individuals, respectively, as variants 3 and 9. Variants 10 and 13, present in 17 and 13% of the population, respectively, contained identical sequences, which differ from 51p1 by seven nucleotides. The remaining five variants occurred in 2–6% of individuals, including two, variants 2 and 11, that encode the same protein as 51p1, and one, variant 4, that encodes the same protein as hv1263. These data demonstrate that, although this allelically related set of V_H1 elements is highly polymorphic, 51p1-containing elements are dominant, and 6 of the 13 elements, with only four different sequences, comprise the alleles present in most people.

Previous studies have revealed that a spectrum exists in the extent and type of polymorphism seen among different V_H loci. The V_H3 gene 20p1 and the single gene of the V_H6 family are nonpolymorphic, in that each is present without apparent sequence variation on one restriction fragment in all individuals studied (8, 19, 22, 25, 39, 40). Such V_H genes have thus undergone little evolutionary alteration in their coding regions or their flanking regions. Other genes, such as the V_H3 element 1.9III (20), several V_H4 elements (21), and the V_H5-2 element (41), virtually always reside on a characteristic restriction fragment but are absent in a subset of the population. The implied null alleles indicate that these V_H genes occupy locations in the V_H cluster that, in some chromosomes, have been deleted. A third type of polymorphism is exemplified by the V_H gene, VH2-4 (41), which always occupies one of two restriction fragments, considered to be an allelic pair. This polymorphism likely represents either a flanking region deletion that spared the V_H coding region or a flanking region point mutation altering a restriction enzyme site. The RFLP we have observed among 51p1, hv1263, and their variants most closely resemble this third type of RFLP, but the variants in the present study reside on a more diverse set of restriction fragments, i.e., six, than has been previously described for V_H alleles.

In most suitably studied V_H loci, coding region sequence variation thus far appears to be either absent (e.g., 20p1 and V_H6) or limited to a few prevalent alleles with nearly identical sequences (8, 19, 20, 22, 25, 39, 40, 42). We identified nine distinct nucleotide sequences in the 51p1 allelic set and confirmed much of their coding region sequence variation by correlation with genomic hybridization results. The data do not suggest that additional sequence variation is common among 51p1 and its alleles. In fact, an exact 51p1 nucleotide sequence was obtained from four different genomes (Table II). It remains unknown, however, if each of the 13 variants, as defined by hybridization, always has the same nucleotide sequence or if new variants would be found with additional subjects. Thus, our data are consistent with the growing evidence that many V_H genes are quite conserved (20, 22) but our data broaden the spectrum of both flanking and coding region polymorphism seen among proven V_H alleles.

The nine nucleotide sequences derived from the 13 hybridization variants have open reading frames, and thus all appear potentially expressible. The encoded amino acid changes that distinguish the sequences include three nonconservative substitutions in the CDR and one in FR3. These differences might be functionally significant because single amino acid changes can significantly alter germline-encoded specificity (43) or specificity for an autoantigen (44). In addition, it is possible that flanking region regulatory sequences differ among the variants, in which case alleles that have the same coding region sequences but reside on different restriction fragments might be differentially expressed.

Mapping of the human V_H cluster indicates that members of the different V_H families are highly interspersed (3, 6, 30) but that the relative chromosomal location of single V_H loci is retained among different haplotypes (27, 29, 30, 45). The fact that 51p1, hv1263, and the related elements are allelically distributed strongly suggests that these variants are all located in the same region of the V_H cluster. We have mapped a 51p1 gene and one of its rare alleles, variants 5 and 8, respectively, in the genome of donor 3116 (Fig. 2, lanes *2*) to similar locations ~ 1,000-kb upstream of V_H6 (30). In addition, a comprehensive map of the V_H genes in the J_H-proximal 800 kb of chromosome 14 did not identify 51p1 or any variants highly homologous to 51p1 (45a). These mapping data are consistent with the conclusion that the 13 variants we have identified probably occupy similar locations on chromosome 14.

Our data indicate that, within the chromosomal region occupied by the 51p1 allelic set, two distinct loci exist in some haplotypes, e.g., the two loci of the 7.5/4.0-kb 51p1-related pair (variants 1 and 7). It is possible that in haplotypes where only one element was identified by our probes a second locus might be present. Undetected V_H segments containing the same CDR as the 13 variants probably do not exist, however, because the CDR probes, which are highly specific, never detected elements other than the 13 we described. It is therefore likely that the single-gene haplotypes do not contain undetected candidate alleles and that the single loci of the 51p1

allelic set evolved either by deleting one segment from a predecessor two-gene haplotype or by never undergoing duplication. In a physical map of the J_H-proximal V_H genes, Shin et al. (27) described a unique V_H1 element present in only one of two otherwise similar haplotypes, indicating that allelic relationships can persist among V_H elements that are adjacent to insertion or deletion polymorphisms. Thus, the most straightforward interpretation of our data is that the 13 variants we have identified comprise an essentially complete set of alleles to a single V_H locus and that in some haplotypes a similar, probably duplicated V_H element is closely adjacent.

Of the allelic sequences we obtained, only 51p1 has been reported in the fetally expressed B cell repertoire (8, 46). Because we found 51p1-containing restriction fragments in 83% of individuals, it is not surprising that the studied fetal genomes had at least one 51p1-containing allele. It does not follow, however, that non-51p1 alleles were absent from the studied fetal genomes because 96% of the individuals we studied had at least one restriction fragment containing a non-51p1 nucleotide sequence. These prevalences reflect the fact that the 51p1-containing 7.5-kb fragment is usually linked to a non-51p1 sequence on a 4.0-kb fragment. Our data thus indicate that the fetal genomes thus far studied likely contained both 51p1 and non-51p1 elements and that fetal expression might favor one or both of the 51p1 alleles (i.e., variants 1 and 5) over non-51p1 alleles. The possibility that the apparent bias toward fetal expression of 51p1 is due to small sample size cannot be excluded.

Several expressed V_H sequences have been reported that are extremely similar to 51p1. Most encode autoantibodies or are expressed by CLL. The Kim13.1 antibody, derived from normal human tonsil, has anticardiolipin specificity and expresses a V_H gene (Humha113) that is identical to our sequence 113T4M27, which we found resides on variant 7, the 4.0-kb fragment of the 7.5/4.0-kb pair (47). It is thus likely that Kim13.1 expresses an unmutated V_H germline gene distinct from 51p1. The V_H sequences AND, NEI, and halLR (renamed from 783c) were expressed by CLL B cell lines (26, 33). AND is identical to 51p1 and NEI and halLR differ only by single nucleotides in the last codon, which could represent somatic junctional modifications of a 51p1 gene. AND, NEI, and halLR probably derived, therefore, from variants 1 or 5, the 51p1 elements we found on 7.5- and 4.3-kb restriction fragments.

The Ab Kim13.1, AND, and NEI carry the V_H1-associated idiotype, G6 (48), which has been found in about one-quarter of CLL (32), one-third of paraprotein IgM RF (49), and 17% of fetal splenocytes (50). Comparison of these expressed V_H sequences with their probable germline sources in the 51p1 allelic set indicates that G6 positivity is associated with the protein sequences of 51p1, which could derive from variants 1, 2, 5, or 11 and at least one allele encoding a non-51p1 protein, variant 7. G6-positive Ab can probably also be encoded by V_H1 genes outside the 51p1 allelic set, because the V_H1 proteins of the G6-positive cryoglobulin IgM RF Sie and Wol differ from 51p1 by more residues than would be expected from somatic substitution alone, i.e., by 38 and 37 amino acids, respectively. The G6 idiotype therefore appears to be encoded by a subset of V_H1 genes that includes, but is not limited to, 51p1 and its alleles.

In conclusion, our data define an allelically related set of V_H1 germline genes whose most prevalent member is the fetally expressed gene 51p1. The flanking and coding region differences that distinguish the alleles could be functionally significant. The actual propensity for each allele to be rearranged or expressed in the normal repertoire, autoimmune states, or B cell neoplasia is not yet known.

Acknowledgments

This work was supported in part by American Cancer Society grant IM485 and National Institutes of Health grants AR-39918, DK-41801, and AR-40561. K. Willems van Dijk was a predoctoral fellow of the March of Dimes Birth Defects Foundation.

References

1. Rechavi, G., B. Bienz, D. Ram, Y. Ben-Neriah, J. B. Cohen, and R. Zakut. 1982. Organization and evolution of immunoglobulin VH gene subgroups. *Proc. Natl. Acad. Sci. USA.* 79:4405–4409.

2. Rechavi, G., D. Ram, L. Glazer, R. Zakut, and D. Givol. 1983. Evolutionary aspects of immunoglobulin heavy chain variable region (VH) gene subgroups. *Proc. Natl. Acad. Sci. USA.* 80:855–859.

3. Kodaira, M., T. Kinashi, I. Umemura, M. Fumihiko, N. Takafumi, Y. Ono, and T. Honjo. 1986. Organization and evolution of variable region genes of the human immunoglobulin heavy chain. *J. Mol. Biol.* 190:529–541.

4. Lee, K. H., F. Matsuda, T. Kinashi, M. Kodaira, and T. Honjo. 1987. A novel family of variable region genes of the human immunoglobulin heavy chain. *J. Mol. Biol.* 195:761–768.

5. Shen, A., C. Humphries, P. Tucker, and F. Blattner. 1987. Human heavy-chain variable region gene family nonrandomly arranged in familial chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA.* 84:8563–8567.

6. Berman, J. E., S. J. Mellis, R. Pollock, C. L. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C. R. Cantor, et al. 1988. Content and organization of the human Ig V_H locus: definition of three new V_H families and linkage to the Ig C_H locus. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:727–738.

7. Humphries, C. G., A. Shen, W. A. Kuziel, J. D. Capra, F. R. Blattner, and P. W. Tucker. 1988. A new human immunoglobulin VH family preferentially rearranged in immature B-cell tumours. *Nature (Lond.).* 331:446–449.

8. Schroeder, H. W., Jr., J. L. Hillson, and R. M. Perlmutter. 1987. Early restriction of the human antibody repertoire. *Science (Wash. DC).* 238:791–793.

9. Mortari, F., J. A. Newton, J. Y. Wang, and H. W. Schroeder, Jr. 1992. The human cord blood antibody repertoire. Frequent usage of the V_H7 gene family. *Eur. J. Immunol.* 22:241–245.

10. Walter, M. A., P. S. Linsley, and D. W. Cox. 1987. Apa I polymorphism of a human immunoglobulin VH3 subclass locus. *Nucleic Acids Res.* 15:4697.

11. Walter, M. A., and D. W. Cox. 1988. Analysis of genetic variation reveals human immunoglobulin VH-region gene organization. *Am. J. Hum. Genet.* 42:446–451.

12. Souroujon, M. C., D. B. Rubenstein, R. S. Schwartz, and K. J. Barrett. 1989. Polymorphisms in human H chain V region genes from the VHIII gene family. *J. Immunol.* 143:706–711.

13. Matthyssens, G., and T. H. Rabbitts. 1980. Structure and multiplicity of genes for the human immunoglobulin heavy chain variable region. *Proc. Natl. Acad. Sci. USA.* 77:6561–6565.

14. Johnson, M. J., A. M. Natali, H. M. Cann, T. Honjo, and L. L. Cavalli-Sforza. 1984. Polymorphisms of a human variable heavy chain gene show linkage with constant heavy chain genes. *Proc. Natl. Acad. Sci. USA.* 81:7840–7844.

15. Turnbull, I. F., O. Bernard, K. Sriprakash, and J. D. Mathews. 1987. Human immunoglobulin variable region genes: a new V_H sequence used to detect polymorphism. *Immunogenetics.* 25:184–192.

16. Pincus, S. H. 1988. Human immunoglobulin heavy-chain-variable (VH) region gene families defined by hybridization with cloned human and murine VH-genes. *Hum. Immunol.* 22:199–215.

17. Chen, P. P., K. A. Siminovitch, N. J. Olsen, R. A. Erger, and D. A. Carson. 1989. A highly informative probe for two polymorphic Vh gene regions that contain one or more autoantibody-associated Vh genes. *J. Clin. Invest.* 84:706–710.

18. Sam, M., M. A. Walter, and D. W. Cox. 1989. RsaI polymorphism of a human immunoglobulin VH5 subclass locus. *Nucleic Acids Res.* 16:8748.

19. Willems van Dijk, K., H. W. Schroeder, Jr., R. M. Perlmutter, and E. C. B. Milner. 1989. Heterogeneity in the human immunoglobulin V_H locus. *J. Immunol.* 142:2547–2554.

20. Sasso, E. H., K. Willems van Dijk, and E. C. B. Milner. 1990. Prevalence and polymorphism of human V_H3 genes. *J. Immunol.* 145:2751–2757.

21. Willems van Dijk, K., E. H. Sasso, and E. C. B. Milner. 1991. Polymorphism of the human immunoglobulin V_H4 gene family. *J. Immunol.* 146:3646–3651.

22. Sanz, I., P. Kelly, C. Williams, S. Scholl, P. Tucker, and J. D. Capra. 1989. The smaller human V_H gene families display remarkably little polymorphism. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:3741–3748.

23. Olee, T., P.-M. Yang, K. A. Siminovitch, N. J. Olsen, J. Hillson, J. Wu, F. Kozin, D. A. Carson, and P. P. Chen. 1991. Molecular basis of an autoantibody-associated RFLP that confers susceptibility to autoimmune diseases. *J. Clin. Invest.* 88:193–203.

24. Chen, P. P., M.-F. Liu, S. Sinha, and D. A. Carson. 1988. A 16/6 idiotype-positive anti-DNA antibody is encoded by a conserved V_H gene with no somatic mutation. *Arthritis Rheum.* 31:1429–1431.

25. Sanz, I., H. Dang, M. Takei, N. Talal, and J. D. Capra. 1989. V_H sequence of a human anti-Sm autoantibody. Evidence that autoantibodies can be unmutated copies of germline genes. *J. Immunol.* 142:883–887.

26. Kipps, T. J., E. Tomhave, L. F. Pratt, S. Duffy, P. P. Chen, and D. A. Carson. 1989. Developmentally restricted immunoglobulin heavy chain variable region gene expressed at high frequency in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA.* 86:5913–5917.

27. Shin, E. K., F. Matsuda, H. Nagaoka, Y. Fukita, T. Imai, K. Yokoyama, E. Soeda, and T. Honjo. 1991. Physical map of the 3′ region of the human immunoglobulin heavy chain locus: clustering of autoantibody-related variable segments in one haplotype. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:3641–3645.

28. Walter, M. A., and D. W. Cox. 1991. Nonuniform linkage disequilibrium within a 1,500-kb region of the human immunoglobulin heavy-chain complex. *Am. J. Hum. Genet.* 49:917–931.

29. Walter, M. A., H. M. Dosch, and D. W. Cox. 1991. A deletion map of the human immunoglobulin heavy chain variable region. *J. Exp. Med.* 174:335–349.

30. Willems van Dijk, K., L. A. Milner, E. H. Sasso, and E. C. B. Milner. 1992. Chromosomal organization of the heavy chain variable region gene segments comprising the human fetal antibody repertoire. *Proc. Natl. Acad. Sci. USA.* 89:10430–10434.

31. Sasso, E. H., K. Willems van Dijk, A. Bull, S. M. van der Maarel, and E. C. B. Milner. 1992. V_H genes in tandem array comprise a repeated germline motif. *J. Immunol.* 149:1230–1236.

32. Kipps, T. J., B. A. Robbins, P. Kuster, and D. A. Carson. 1988. Autoantibody-associated cross-reactive idiotypes expressed at high frequency in chronic lymphocytic leukemia relative to B-cell lymphomas of follicular center cell origin. *Blood.* 72:422–428.

33. Chen, P. P., M.-F. Liu, C. A. Glass, S. Sinha, T. J. Kipps, and D. A. Carson. 1989. Characterization of two immunoglobulin V_H genes that are homologous to human rheumatoid factors. *Arthritis Rheum.* 32:72–76.

34. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

35. Kabat, E. A., T. T. Wu, H. M. Perry, K. S. Gottesman, and C. Foeller. 1991. Sequences of Proteins of Immunological Interest. U.S. Department of Health and Human Services, U.S. Public Health Service, NIH, U.S. Government Printing Office, Bethesda, MD.

36. Wood, W. I., J. Gitschier, L. A. Lasky, and R. M. Lawn. 1985. Base composition-independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA.* 82:1585–1588.

37. Dunning, A. M., P. Talmud, and S. E. Humphries. 1988. Errors in the polymerase chain reaction. *Nucleic Acids Res.* 16:10393.

38. Keohavong, P., and W. G. Thilly. 1989. Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. USA.* 86:9253–9257.

39. Schroeder, H. W., Jr., M. A. Walter, M. H. Ho f ker, A. Ebens, K. Willems van Dijk, L. C. Liao, D. Cox, E. C. B. Milner, and R. M. Perlmutter. 1988. Physical linkage of a human immunoglobulin heavy chain variable region gene segment to diversity and joining region elements. *Proc. Natl. Acad. Sci. USA.* 85:8196–8200.

40. Logtenberg, T., F. M. Young, J. H. Van Es, F. H. J. Gmelig-Meyling, and F. W. Alt. 1989. Autoantibodies encoded by the most J_H-proximal human immunoglobulin heavy chain variable region gene. *J. Exp. Med.* 170:1347–1355.

41. Walter, M. A., W. T. Gibson, G. C. Ebers, and D. W. Cox. 1991. Susceptibility to multiple sclerosis is associated with the proximal immunoglobulin heavy chain variable region. *J. Clin. Invest.* 87:1266–1273.

42. Buluwela, L., and T. H. Rabbitts. 1988. A V_H gene is located within 95 kb of the human immunoglobulin heavy chain constant region genes. *Eur. J. Immunol.* 18:1843–1845.

43. Sikder, S. K., P. Borden, F. Gruezo, P. N. Akolkar, S. B. Bhattacharya, S. L. Morrison, and E. A. Kabat. 1989. Amino acid substitutions in V_H CDR2 change the idiotype but not the antigen-binding of monoclonal antibodies to $\alpha(1\rightarrow6)$dextrans. *J. Immunol.* 142:888–893.

44. Diamond, B., and M. D. Scharff. 1984. Somatic mutation of the T15 heavy chain gives rise to an antibody with autoantibody specificity. *Proc. Natl. Acad. Sci. USA.* 81:5841–5844.

45. Willems van Dijk, K., L. A. Milner, and E. C. B. Milner. 1992. Mapping of human H chain V region genes (V_H4) using deletional analysis and pulsed field gel electrophoresis. *J. Immunol.* 148:2923–2931.

45a. Matsuda, F., E. K. Shin, H. Nagaoka, R. Matsumura, M. Haino, Y. Fukita, S. Taka-ishi, T. Imai, J. H. Riley, R. Anand, E. Soeda, and T. Honjo. 1993. Structure and physical map of 64 variable segments in the 3′ 0.8-megabase region of the human immunoglobulin heavy-chain locus. *Nature Genetics.* 3:88–94.

46. Schroeder, H. W., Jr., and J. Y. Wang. 1990. Preferential utilization of conserved immunoglobulin heavy chain variable gene segments during human fetal life. *Proc. Natl. Acad. Sci. USA.* 87:6146–6150.

47. Siminovitch, K. A., V. Misener, P. C. Kwong, P.-M. Yang, C. A. Laskin, E. Cairns, D. Bell, L. A. Rubin, and P. P. Chen. 1990. A human anti-cardiolipin autoantibody is encoded by developmentally restricted heavy and light chain variable region genes. *Autoimmunity.* 8:97–105.

48. Mageed, R. A., M. Dearlove, D. M. Goodall, and R. Jefferis. 1986. Immunogenic and antigenic epitopes of immunoglobulins XVII—Monoclonal antibodies reactive with common and restricted idiotopes to the heavy chain of human rheumatoid factors. *Rheumatol. Int.* 6:179–183.

49. Silverman, G. J., R. D. Goldfien, P. Chen, R. Mageed, R. Jefferis, F. Goni, B. Frangione, S. Fong, and D. A. Carson. 1988. Idiotypic and subgroup analysis of human monoclonal rheumatoid factors. Implications for structural and genetic basis of autoantibodies in humans. *J. Clin. Invest.* 82:469–475.

50. Kipps, T. J., B. A. Robbins, and D. A. Carson. 1990. Uniform high frequency expression of autoantibody-associated crossreactive idiotypes in the primary B cell follicles of human fetal spleen. *J. Exp. Med.* 171:189–196.