IDENTIFICATION AND CHARACTERIZATION OF NEW MURINE T CELL RECEPTOR β CHAIN VARIABLE REGION (V_β) GENES

By MARJORIE C. LOUIE, CHRISTOPHER A. NELSON, and DENNIS Y. LOH

From the Departments of Medicine, Genetics, and Microbiology, and Immunology, Howard Hughes Medical Institute, Washington University School of Medicine, St. Louis, Missouri 63110

The antigen receptor that is present on the surface of most T cells is a heterodimer composed of disulfide-linked α and β chains, each of which has C and V regions (see reference 1 for a review). The V region of each β chain is encoded by noncontiguous variable (V_β), diversity (D_β), and joining (J_β) gene segments that are juxtaposed during T cell development by the process of somatic rearrangement. To date, only 22 murine TCR β chain V region (V_β) gene segments (20 functional genes and two V_β pseudogenes) have been identified, largely from analysis of C_β^+ clones from cDNA libraries (1–3). A statistical analysis of the first 25 V_β sequences identified (representing 14 independent V_β gene segments) yielded a maximum likelihood estimate of 18 for the maximum number of V_β genes, with a 95% one-sided confidence bound of 30 (4). This estimate, however, is based on the assumption that each of the V_β genes is equally likely to be used, whereas actual gene usage may be nonrandom. In fact, the actual V_β repertoire is affected by at least two mechanisms, deletion and nonexpression. Examples of gene deletion are the V_β^a haplotype strains of mice, such as SJL, SWR, C57BR, C57L, and AU/SsJ, which have a deletion within the V_β locus that eliminates 10 contiguous V_β gene segments (5–7). Examples of the nonexpression of a V_β gene are the instances of clonal elimination of T cells bearing α/β receptors that are reactive to so-called "superantigens" in combination with certain H-2 class II molecules (8–14).

If, in fact, V_β gene usage is nonrandom, then there may exist many V_β genes that are used at low frequency and have not been isolated from cDNA libraries. However, only one complete new functional V_β gene, V_β17a (15), has been identified in the last 3 yr. Isolation of another V_β gene, provisionally called V_β18, has been reported, but only a partial nucleotide sequence of this gene has been published (16). We sought to investigate whether there exist yet-undiscovered V_β genes that are expressed at low frequency and whether there exists a substantial number of pseudogenes by using a panel of cosmid clones, spanning the BALB/c V_β gene locus, that had been previously isolated in our laboratory (6).

In this paper, we report a novel method for identifying new V_β genes and present details on seven new V_β genes. The new method involves probing cloned genomic

Address correspondence to Dennis Y. Loh, Howard Hughes Medical Institute, Washington University School of Medicine, 660 South Euclid Avenue, Box 8045, St. Louis, MO 63110.

J. Exp. Med. © The Rockefeller University Press · 0022-1007/89/12/1987/12 $2.00 1987
Volume 170 December 1989 1987–1998

DNA with a Vβ-specific oligonucleotide instead of screening cDNA with a Cβ-specific probe, thereby circumventing the problem, encountered with screening cDNA, of low expression of certain functional Vβ genes. We report here the discovery of a new functional Vβ gene, Vβ19a, and six pseudogenes, including Vβ19b, as well as further studies on the expression of Vβ19. In addition, we present the complete nucleotide sequence of BALB/c Vβ18, which was isolated from cloned genomic DNA after probing with a cDNA probe.

Materials and Methods

Mice. 5-wk-old BALB/cJ, C57BL/6J, SJL/J, SWR/J, C57BR/cdJ, and C57L/J mice were obtained from The Jackson Laboratory, Bar Harbor, ME.

Oligonucleotides. Oligonucleotides, which were used as hybridization probes, sequencing primers, and polymerase chain reaction (PCR)[1] primers, were synthesized on a DNA synthesizer (380A; Applied Biosystems, Inc., Foster City, CA).

Screening of Cosmids with a Vβ-specific Oligonucleotide. The overlapping cosmid clones spanning 330 kb of the Vβ locus in BALB/c, along with the restriction map, have been previously described (6).

Cosmid DNA was digested with restriction endonucleases and separated on a 0.7% agarose gel. The gel was dried, denatured in 0.5 N NaOH, 1.5 M NaCl, neutralized in 0.5 M Tris (pH 7.6), 0.75 M NaCl, and hybridized at 42°C with a Vβ-specific oligonucleotide, called the CASS oligonucleotide (described in Results and shown in Fig. 1), that had been 5′ end labeled with T4 polynucleotide kinase and γ-[^{32}P]ATP. The hybridization solution consisted of 0.9 M NaCl, 0.18 M Tris (pH 7.4), 12 mM EDTA, 10 mM sodium pyrophosphate, and 2× Denhardt's solution. After hybridization, gels were washed at room temperature with 5× SSC (0.75 M NaCl, 0.075 M sodium citrate) and then exposed to film, with an intensifying screen, at −70°C.

Six CASS oligonucleotide-positive restriction fragments that did not contain previously reported Vβ genes were identified (designated by an N, for new, followed by a number) and subcloned into pBluescribe (Stratagene, La Jolla, CA) to generate the corresponding plasmids (pN1, pN2, etc.).

DNA Sequencing. The sequence of potential Vβ genes in the subcloned CASS oligonucleotide-positive restriction fragments was determined by the chain-termination method (17), as modified for sequencing double-stranded plasmid DNA with Sequenase DNA polymerase (United States Biochemical Corp., Cleveland, OH). Initial sequence data were obtained using the CASS oligonucleotide as sequencing primer; additional sequence data upstream, downstream, and on the opposite strand were obtained using sequencing primers derived from the initial sequence.

Vβ18 had been previously localized on the genomic map (6) by hybridization of restricted cosmid DNA with a Vβ18 cDNA probe provided by Dr. P. A. Singer (Scripps Clinic and Research Foundation, La Jolla, CA). For sequencing, a 4-kb Bam HI-Sal I fragment from cosmid C55 was subcloned into pBluescribe, and the nucleotide sequence of Vβ18 was obtained by the method of Maxam and Gilbert (18).

Sequence analyses were performed using the Microgenie sequence analysis program (Beckman Instruments, Inc., Palo Alto, CA).

RNase Protection. Total cellular RNA from spleen and thymus of 5-wk-old mice of six different strains was prepared by the guanidinium isothiocyanate method (19).

RNA probes were synthesized according to the protocol from the supplier (Stratagene). Antisense Vβ16 RNA was synthesized from a DNA template that consisted of a genomic Xba I fragment containing 67 bp of the intron between the first and second Vβ exons, 293 bp of the second exon, and ~220 bp of 3′ flanking region. Antisense Vβ19 RNA was synthesized from a DNA template that consisted of a genomic Nhe I fragment containing 79 bp

[1] *Abbreviation used in this paper: PCR, polymerase chain reaction.*

of leader exon and adjacent 5′ untranslated region, 112 bp of intron between the first and second Vβ exons, and 285 bp of second exon.

RNase protection was performed according to the method of Melton et al. (20), with minor modifications as detailed in Anderson et al. (21).

Southern Blot Analysis. 15 μg of high molecular weight DNA from BALB/c and SJL liver was digested with restriction endonucleases and separated on a 0.7% agarose gel. The gel was dried, denatured, and neutralized as described for the screening of cosmids, and then hybridized with a uniformly ^{32}P-labeled probe generated by random hexanucleotide-primed synthesis by the Klenow fragment of *Escherichia coli* DNA polymerase I in the presence of α-[^{32}P]dCTP (22).

The VβN3 probe was a 285-bp Nhe I–Nco I fragment from pN3, containing almost the entire second exon (which encodes a few amino acids of the leader peptide and all amino acids of the mature Vβ polypeptide segment). The VβN1 probe was a 410-bp Bsm I–Hpa I fragment from pN1, containing the entire second exon, ~40 bp of upstream intron, and ~90 bp of 3′ flanking sequence.

Gels were hybridized at 65°C in 5× SSC, 5× Denhardt's solution, 10 mM sodium pyrophosphate, and 100 μg/ml sheared, denatured salmon sperm DNA. After hybridization, gels were washed with 2× SSC (0.3 M NaCl, 0.03 M sodium citrate) at 65°C and then exposed to film, with an intensifying screen, at −70°C.

Cloning of Vβ19 from SJL by the PCR. PCR was performed according to the protocol supplied by Perkin-Elmer Cetus (Norwalk, CT). 1 μg of high molecular weight DNA from SJL liver was subjected to 30 cycles of PCR in a total volume of 100 μl, containing 50 mM KCl, 10 mM Tris (pH 8.3), 1.5 mM $MgCl_2$, 1.5 mM of each dNTP, gelatin at 100 μg/ml, ~1 μg of each primer, and 5 U of Taq polymerase (Perkin-Elmer Cetus). Except for the first cycle, which included a prolonged 5-min denaturation step, and the last cycle, which included a prolonged 5-min polymerization step, the intervening 28 PCR cycles consisted of denaturation at 90°C for 1.5 min, annealing at 55°C for 1 min, and polymerization at 70°C for 2 min. The PCR product was cloned into the Sma I site of pBluescribe, and plasmid DNA was sequenced as described above.

Initially, a 5′ primer (PCR N3-5a), containing sequence upstream of the BALB/c Vβ19 leader, and a 3′ primer (PCR N3-3a), containing sequence close to the 3′ end of the second exon, were used to obtain an amplified fragment including the first (leader) exon and most of the second exon; three independent clones containing this fragment were sequenced. Subsequently, another 5′ primer (PCR N3-5b), containing sequence upstream of the leader, and another 3′ primer (PCR N3-3d), containing sequence in the 3′ untranslated region, were used to obtain a larger amplified fragment containing all of both Vβ exons; one clone containing this fragment was sequenced. The sequences of the primers are as follows: PCR N3-5a: 5′-GAAGAATTCATTGTGTATGAGAAGGTAGCCTAAG-3′; PCR N3-3a:5′-AGAGAATTC-ATCTGTTAGCTTGCTGGACTGGATC-3′; PCR N3-5b: 5′-GAAGAATTCACCTTCAG-AGAAGACAGCTAGCTGA-3′; PCR N3-3d: 5′-AGAGAATTCGCTGATCAGCACTCT-CTTGCTTCTT-3′.

Results

Design of a Vβ-specific Oligonucleotide. To design a Vβ-specific oligonucleotide that could be used to screen cloned genomic DNA for new Vβ genes, we examined the amino acid sequences encoded by the known Vβ genes for amino acid stretches that are highly conserved. It was noted that 14 of the 19 Vβ genes that had been identified by the beginning of this study encoded Tyr-Phe/Leu-Cys-Ala-Ser-Ser as the last six COOH-terminal amino acids (Fig. 1). From an examination of the nucleotide sequence that encodes these six amino acids, we arrived at a fourfold degenerate 18-mer (CASS oligonucleotide) that is the inverse complement of the coding sequence (Fig. 1).

Identification of New Vβ Genes. Southern analysis of restricted DNA from a series

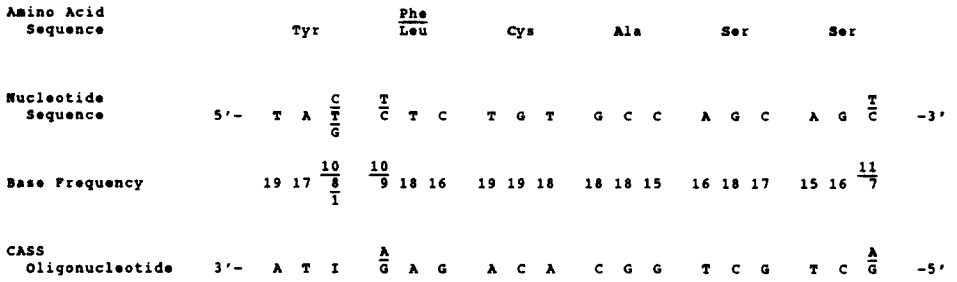

FIGURE 1. Derivation of the nucleotide sequence of the CASS oligonucleotide. (Line *1*) The amino acid sequence encoded by the 3' end of 14 of 19 Vβ gene segments. (Line *2*) The nucleotide sequence encoding the above six amino acids. (Line *3*) The frequency of base usage, tabulated from an examination of 19 different Vβ sequences, at each of the 18 nucleotide positions shown in line *2*. (Line *4*) The sequence of the CASS oligonucleotide; the inverse complement of the sequence shown in line *2*.

of cosmid clones spanning the Vβ locus (including a series of overlapping cosmids encompassing the 330-kb region containing 20 of the previously identified Vβ genes, two overlapping cosmids containing Vβ2, and two contiguous cosmids containing Cβ1, Cβ2, and Vβ14) yielded six CASS oligonucleotide-positive restriction fragments (designated by the prefix N in Fig. 2) that did not contain one of the known Vβ genes. A detailed map showing restriction sites and cosmids has been previously published (6).

Two of the CASS oligonucleotide-positive fragments, N8 and N9, are located in the 55-kb gap between Vβ1 and Vβ5.2. Two other fragments, N3 and N1, are located upstream and downstream, respectively, of Vβ17b.

Nucleotide and Predicted Amino Acid Sequences of the New Vβ Genes. Two criteria that were used to assess whether a sequence is that of a Vβ gene are the presence of heptamer-nonamer recombination signals downstream of the second exon and the presence, in the predicted amino acid sequence, of the six amino acids that are absolutely conserved in all Vβ genes discovered to date.

The nucleotide sequences of putative Vβ genes in the CASS oligonucleotide-positive fragments are shown in Fig. 3. Five of the six new CASS oligonucleotide-

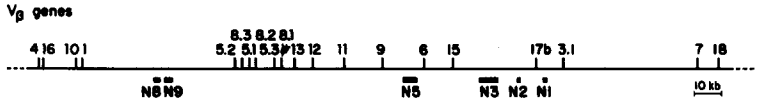

FIGURE 2. Map of a portion of the BALB/c Vβ locus, showing the location of six CASS oligonucleotide-positive restriction fragments that do not contain known Vβ genes. The Vβ genes are indicated by vertical lines above the horizontal line. CASS oligonucleotide-positive restriction fragments not containing known Vβ genes are depicted by solid bars under the horizontal line and are designated by N (for new), followed by a number. A detailed map showing the restriction sites and the cosmid clones spanning this region of the genome has been previously published (6). N1 is a 1.5-kb Kpn I-Eco RI fragment contained within the 8-kb Bam HI fragment of cosmid D18. N2 is a 1.1-kb Bam HI fragment from cosmid D18. N3 is a 7-kb Kpn I fragment from cosmid B2, which has an insert that is almost the same as that of cosmid B47. N5 is a 6-kb Kpn I-Sal I fragment from cosmid B47. N8 is a 2.4-kb Kpn I fragment from cosmid B41. N9 is a 4-kb Kpn I-Bam HI fragment from cosmid C47.

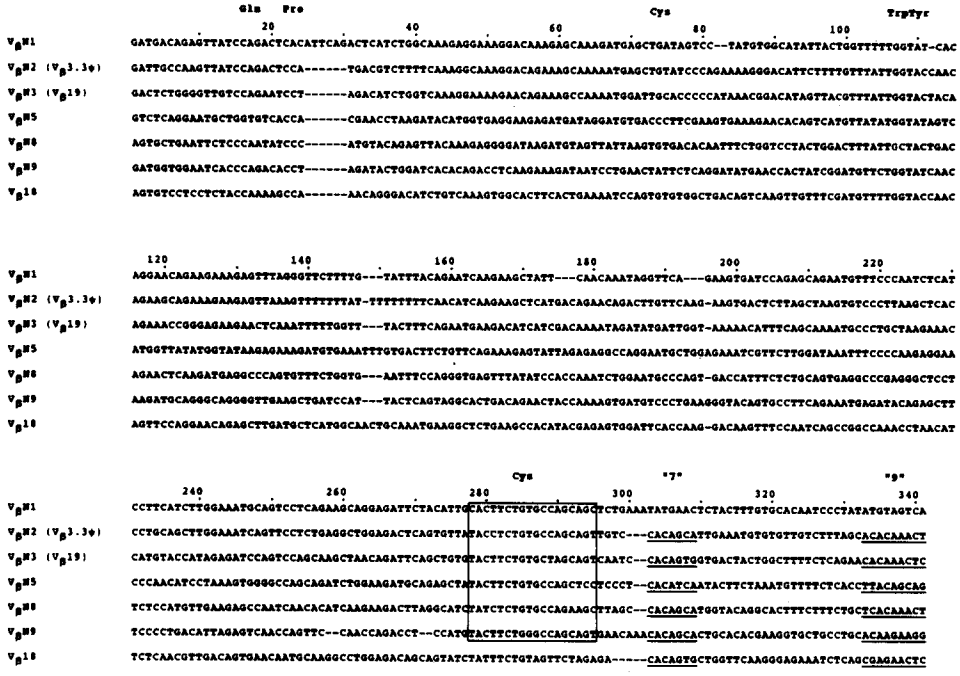

FIGURE 3. Nucleotide sequence of the putative new Vβ genes in the six CASS oligonucleotide-positive restriction fragments. The putative new Vβ genes are named after the restriction fragments in which they lie. Initial sequence was obtained using the CASS oligonucleotide as sequencing primer; additional sequence upstream, downstream, and on the opposite strand was obtained with oligonucleotide primers derived from the initial sequence. The various sequences were aligned by first aligning the nucleotides encoding the invariant amino acids, which are shown (*boldfaced*) above the nucleotide sequence. Nucleotides are numbered according to Patten et al. (23), with position 1 being the first base of the codon specifying the first amino acid of the mature Vβ polypeptide segment after cleavage of the leader peptide. The sequence that hybridized with the CASS oligonucleotide is boxed. Putative heptamer and nonamer recombination signals downstream of the protein-coding region are underlined and denoted by *7* and *9*, respectively. Dashes within the sequence indicate gaps that have been introduced to maximize homology. Shown in the last line is the sequence of Vβ18, a gene for which a partial nucleotide sequence has been previously published (16). These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession numbers X16689-X16695.

positive genes contain heptamer and nonamer recombination signals downstream of the protein-coding sequence. The exception, VβN1, has a heptamer-like sequence, CACAATC, somewhat farther downstream than the heptamer found in other Vβ genes, but no nonamer-like sequence 23 bp downstream. The transcriptional orientation of all six new CASS oligonucleotide-positive genes is the same as that of previously identified Vβ genes in the contiguous 330-kb genomic region screened.

The amino acid sequences (Fig. 4) predicted by the nucleotide sequences were also examined to determine whether the encoded amino acids were characteristic of a Vβ polypeptide segment. Three of the six new Vβ genes (VβN3, VβN8, and VβN9) have identifiable invariant amino acids in a single reading frame; thus, the amino acid sequence presented in Fig. 4 is that predicted by the nucleotide sequence of the single reading frame. The other three new CASS oligonucleotide-positive Vβ

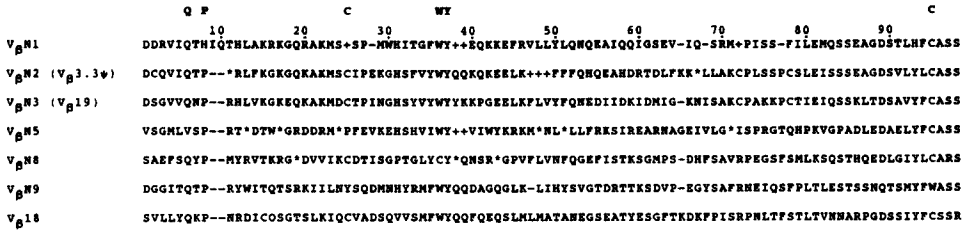

FIGURE 4. Amino acid sequences of the six putative new CASS oligonucleotide-positive Vβ genes. The amino acids are denoted by the single-letter code. The various amino acid sequences are aligned by the six amino acids that are invariant in the Vβ genes identified to date; these six invariant amino acids are shown (*boldfaced*) above the sequences. For those Vβ genes (VβN3, VβN8, and VβN9) that have the invariant amino acids in a single reading frame, the amino acid sequence shown is that predicted by the nucleotide sequence of the single reading frame. For those Vβ genes (VβN1, VβN2, and VβN5) that have the invariant amino acids in multiple reading frames, the sequence shown is a "best-fit" amino acid sequence; the transition region between two reading frames, where a definite amino acid sequence cannot be deduced, is indicated by one or more plus signs. A dash within a sequence indicates a gap that has been introduced to optimize alignment. An asterisk indicates the position corresponding to a termination codon.

genes (VβN1, VβN2, and VβN5) have identifiable invariant amino acids in multiple reading frames, presumably due to nucleotide insertions and/or deletions; therefore, the amino acid sequence presented in Fig. 4 is a "best-fit" sequence, with a plus sign indicating the transition between two reading frames, where a specific amino acid sequence cannot be deduced.

VβN1 and VβN5 contain duplicated nucleotide stretches that disrupt the reading frame. The duplication in VβN1 includes the proline residue at position 8 that is absolutely conserved in all previously reported Vβ genes; in addition, VβN1 is missing one of the two invariant cysteine residues. The duplication in VβN5 results in a duplication of the conserved tryptophan and tyrosine residues at positions 36 and 37; VβN5 is also missing one of the two invariant cysteines and contains at least six termination codons. VβN8 contains three termination codons and is missing the conserved tryptophan residue at position 36. VβN2, which contains at least two termination codons, is identical to a pseudogene, Vβ3.3ψ, which had been identified earlier in our laboratory by crosshybridization with a Vβ3 probe.

Two of the new oligonucleotide-positive genes, VβN3 and VβN9, have open reading frames. VβN9, however, contains neither of the conserved cysteine residues that are presumably involved in the intrachain disulfide bond. Thus, on the basis of sequence alone, five of the six new CASS oligonucleotide-positive genes cannot encode functional Vβ polypeptide segments and are presumably pseudogenes. On the other hand, the open reading frame of VβN3 encodes all six of the amino acids that are invariant in the previously reported Vβ genes; this potentially functional gene we call Vβ19.

Expression of Vβ19. Because it had an open reading frame that encoded all six of the amino acids absolutely conserved in functional Vβ genes, and because it had heptamer-nonamer recombination signals downstream of the protein-coding region, Vβ19 was potentially a functional gene and was therefore further characterized. To determine whether Vβ19 is expressed, we performed RNase protection experiments using total cellular RNA and an RNA probe encompassing both exons of Vβ19 (in-

cluding sequence upstream of the leader exon) and the intervening intron. These experiments showed that Vβ19 was not transcribed in the thymus of BALB/c mice, the strain from which it was identified, but was transcribed in SJL thymus (Fig. 5 *A*). SJL is a V_β^a haplotype strain, in which a large deletion has eliminated 10 contiguous Vβ gene segments, and in which nondeleted Vβ genes might be compensatorily expressed at increased frequency. RNase protection analysis of RNA from another V_β^b haplotype (nondeletion) strain (C57BL/6) and from three other V_β^a haplotype strains (SWR, C57BR, and C57L) confirms the pattern of expression found

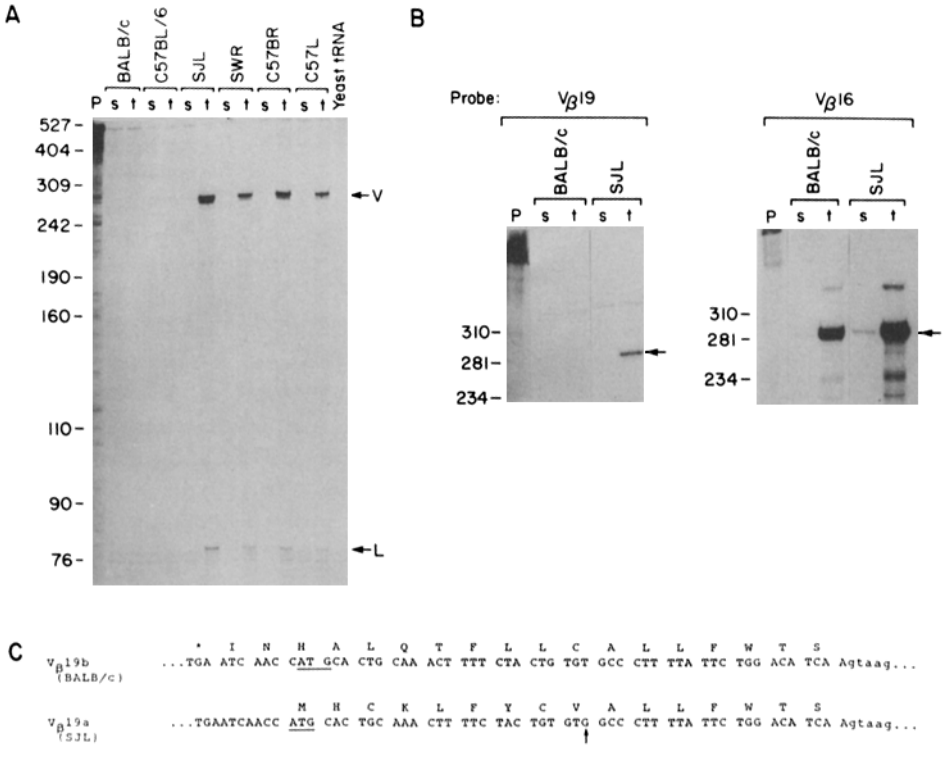

FIGURE 5. RNase protection analysis of Vβ19 (VβN3). (*A*) Autoradiograph of a gel from RNase protection assays using a Vβ19 probe against RNA samples from two V_β^b haplotype strains and from four V_β^a haplotype strains of mice. 50 μg of total cellular RNA was used in each assay. Details about the antisense RNA probe are given in Materials and Methods. As a negative control, the Vβ19 probe was assayed against yeast tRNA. The protected fragments are indicated by arrows labeled with V (for the Vβ second exon) and L (for the leader exon and 5′ flanking sequences). P indicates a lane in which only RNA probe is loaded. s, spleen; t, thymus. (*B*) Level of expression of Vβ19 compared with that of Vβ16. As in *A*, 50 μg of total cellular RNA was used in each assay. Details about the Vβ19 and Vβ16 antisense RNA probes are given in Materials and Methods. The protected fragments, corresponding to the Vβ second exons, are indicated by arrows on the right side of each panel. The sizes, in nucleotides, of selected fragments of an Msp I digest of pBR322 are shown on the left side of each panel. P, probe only; s, spleen; t, thymus. (*C*) Comparison of Vβ19 leader sequences from BALB/c and from SJL. ATG (initiation) codons, out-of-frame in BALB/c and in-frame in SJL, are underlined. The single-base insertion in the SJL sequence that results in an in-frame ATG upstream is indicated by an arrow. The lowercase letters at the 3′ end of the sequence are the nucleotides at the 5′ end of the intron between the first (leader) and second Vβ exons.

in BALB/c and SJL; that is, absent (or undetectably low) expression in V_β^b haplotype strains and low-level expression in V_β^a haplotype strains (Fig. 5 *A*). Comparison with Vβ16, a Vβ gene expressed in ~5% of Cβ-containing clones from an SJL thymus cDNA library (M. C. Louie, unpublished data) shows that the level of expression of Vβ19 in SJL thymus is low, considerably less than that of Vβ16 (Fig. 5 *B*).

RNase protection analysis of RNA from thymus of V_β^a haplotype mice also reveals a protected fragment of ~80 nucleotides (Fig. 5 *A*), consistent with protection of the portion of the RNA probe corresponding to the leader exon and the adjacent upstream sequence of BALB/c Vβ19b. Thus, in the V_β^a haplotype strains of mice, a leader sequence that is identical to, or nearly identical to, the BALB/c Vβ19b leader is transcribed.

To investigate whether a difference in the sequences of the Vβ19 genes from V_β^a and V_β^b haplotype mouse strains may account for the difference in expression of the gene, we cloned the Vβ19 gene from SJL genomic DNA by the PCR method, using oligonucleotide primers corresponding to sequences upstream of the BALB/c Vβ19 leader sequence and within or downstream of the BALB/c Vβ19 second exon. Comparison of the nucleotide sequences of the Vβ19 genes from BALB/c and from SJL revealed that the exon sequences are identical, except for a single-base insertion in the leader exon of Vβ19 from SJL, shifting the reading frame so that there is no longer an in-frame termination codon downstream of an out-of-frame initiation codon, as was found in the leader exon of Vβ19 from BALB/c, thereby resulting in a leader exon encoding 19 amino acids (Fig. 5 *C*). Thus, in SJL, Vβ19 is a functional gene.

Southern Blot Analysis of Vβ19. Southern analysis of genomic DNA from BALB/c and from SJL showed RFLP at the Vβ19 locus for five different restriction endonucleases (Fig. 6). Similarly, the VβN1 locus in BALB/c and in SJL showed RFLP for Eco RI, Hind III, and Bgl II (data not shown). Genomic DNA prepared from the tails of mice of the six strains used in the RNase protection experiments (Fig. 5 *A*) and hybridized with the Vβ19 second exon probe showed two-allele RFLP for

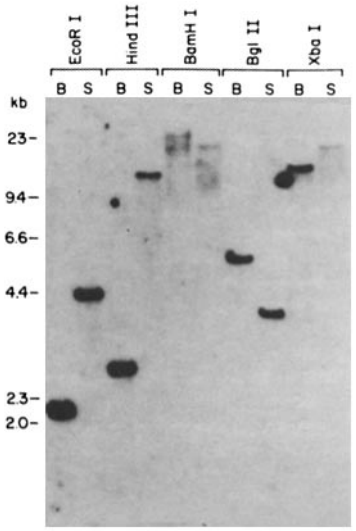

FIGURE 6. Southern blot analysis of total genomic DNA from BALB/c and from SJL. 15 μg of high molecular weight DNA prepared from BALB/c or SJL liver was digested with the restriction enzymes indicated, and the DNA was separated on a 0.7% agarose gel. The dried and denatured gel was hybridized with a ^{32}P-labeled Vβ19 (VβN3) probe; hybridization and wash conditions are given in Materials and Methods. Size markers, in kilobases, are shown at the left. B, BALB/c; S, SJL.

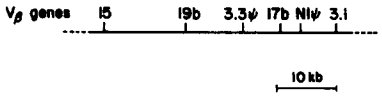

FIGURE 7. Map of the 39-kb region of DNA bounded by Vβ15 and Vβ3.1 that contains two functional Vβ genes and four Vβ pseudogenes. The transcriptional orientation of all six Vβ genes is the same, with the 5' to 3' direction running left to right on the map. A restriction map of this region has been previously published (6).

both Kpn I (7 kb in the V_β^b haplotype mice; 9 kb in the V_β^a haplotype mice) and Eco RI (2.1 kb in the V_β^b haplotype mice; 4.3 kb in the V_β^a haplotype mice) (data not shown). A map of this genomic region is shown in Fig. 7.

Discussion

In this paper, we report the nucleotide and amino acid sequences of six new Vβ genes that were isolated by a novel method in which cloned genomic DNA is probed with a fourfold degenerate oligonucleotide that is complementary to the nucleotide sequence encoding six amino acids that are present at the COOH terminal of approximately three fourths of the known Vβ polypeptide segments. Probing genomic DNA circumvents the difficulty of cloning genes that are expressed at very low frequency in a cDNA library. Since only about three fourths of the previously identified functional Vβ genes encode a polypeptide segment with Tyr-Phe/Leu-Cys-Ala-Ser-Ser at the COOH-terminal end, we might expect, by extrapolation, that probing with the CASS oligonucleotide would fail to identify about one fourth of the yet-unidentified Vβ genes.

Five of the six CASS oligonucleotide-positive genes are pseudogenes, based on the presence of termination codons within the coding region or the absence of absolutely conserved amino acids, especially one or both of the cysteine residues that form the intrachain disulfide bond. Several of the new genes contain obvious frameshifts, suggesting the insertion or deletion of one or more nucleotides. Nevertheless, based on the presence of other absolutely conserved amino acids and on the presence of heptamer and nonamer recombination signal sequences downstream of the protein-coding sequence, the newly identified genes are clearly recognizable as Vβ gene segments.

Because only two of the first 22 Vβ genes were pseudogenes, it was concluded that Vβ pseudogenes are relatively rare. In contrast, the frequency of pseudogenes among Ig V_H genes is ~30% (24). Here, however, we report the discovery of six new Vβ pseudogenes, indicating that Vβ pseudogenes are much more abundant than previously thought. An examination of the nucleotide sequences of these six pseudogenes reveals that the present nucleotide sequences can be derived from sequences encoding all six of the amino acids that are absolutely conserved among known Vβ gene segments by simple (often single-nucleotide) substitutions, insertions, or deletions. The pseudogenes that have been identified to date fall into two classes: pseudogenes (such as Vβ17b and Vβ19b) with point mutations, and pseudogenes with more extensive mutations involving multiple regions of the coding sequences. The observation that all six of the new CASS oligonucleotide-positive Vβ genes in BALB/c are pseudogenes is more consistent with the hypothesis that the functional Vβ repertoire is small and that most of the functional Vβ genes have already been cloned than with the notion that the functional Vβ repertoire is large, with many Vβ genes used infrequently.

Two of the new pseudogenes, VβN1 and VβN2 (Vβ3.3ψ), are located in the relatively crowded genomic region bounded by Vβ15 and Vβ3.1, where two functional Vβ gene segments and four Vβ pseudogenes are located within 39 kb of DNA (Fig. 7). Further analysis of the nucleotide sequences (Fig. 3) reveals that these clustered genes are highly homologous. Not unexpectedly, the second and third members of the Vβ3 subfamily (Vβ17b and Vβ3.3ψ) are both 74% identical, at the nucleotide level, to Vβ3.1. VβN1 is slightly less homologous (66–68% identical at the nucleotide level) to the three members of the Vβ3 subfamily. Vβ19b is even less homologous (64–65% identical at the nucleotide level) to the Vβ3 subfamily members and only 57% identical to VβN1. The extensive homology among these genes makes possible the process of recombination by homologous but unequal crossing over. Though the molecular mechanism and the specific sequence requirements of the recombination process are largely unknown, the extensive RFLP noted for Vβ19 and VβN1 (this paper), and for Vβ17 (15), supports the notion that this region of highly homologous pseudogenes is the site of an increased rate of recombination. Interestingly, a newly described Vβ deletion has its 3′ boundary in this region (25).

The RFLP at the Vβ19 locus in the V_β^a and V_β^b haplotype mouse strains is analogous to that previously reported for Vβ17 (15). In keeping with the nomenclature established with Vβ17 (3, 15), we call Vβ19a the allele that is present in V_β^a haplotype mouse strains and Vβ19b the allele that is present in V_β^b haplotype strains. RNase protection experiments reveal that a leader sequence identical to, or nearly identical to, the leader sequence from BALB/c, a V_β^b haplotype strain, is indeed transcribed (and presumably correctly spliced to a Vβ19a second exon) in the V_β^a haplotype mice. Thus, the expression of Vβ19a does not appear to involve alternative splicing of the transcripts from the second exon of Vβ19a and from the intact leader exon of another Vβ gene, as has been reported with the alternative splicing of the transcripts from the second exon of Vβ8.2 and from the leader exon of Vβ5.1, a gene located ~2.5 kb upstream of Vβ8.2 (2).

It has been shown that T lymphocytes expressing Vβ17a are clonally eliminated in mouse strains, such as C57BR, that express an I-E class II MHC molecule (5, 14, 15). New gene Vβ19 was expressed in all four V_β^a haplotype strains examined and not expressed in both V_β^b haplotype strains examined. By analogy, if Vβ19a-carrying strains that do not express Vβ19a can be found, one might investigate whether Vβ19a is clonally eliminated because of reactivity to a self-antigen expressed in these strains.

Summary

By screening previously isolated genomic clones spanning the mouse TCR Vβ locus with Vβ-specific oligonucleotides, we have isolated one new functional Vβ gene and six Vβ pseudogenes. Because this method of identifying new genes does not depend on expression levels, we conclude that most, if not all, Vβ genes in the mouse have been identified. The newly identified pseudogenes increase the frequency of mouse TCR Vβ pseudogenes to 28%, a frequency similar to that estimated for mouse Ig V_H pseudogenes (24).

Three of the newly discovered pseudogenes are clustered in a region around another pseudogene (Vβ17b). The extensive DNA diversity, as reflected in both the

nucleotide sequence and the RFLP, indicates that this genomic region is a possible hotspot of recombination.

The new functional gene, Vβ19a, is expressed at very low levels, which explains why it has not been isolated earlier. Vβ19 shows expression patterns that correlate with the previously described V_β^a and V_β^b haplotypes.

We thank Dr. P. A. Singer for the Vβ18 cDNA probe; Dr. G. Dugan for the BALB/c and SJL genomic DNA that was used in the Southern blot analyses and the PCR reactions; Mr. A. Cheng for assistance with the PCR technique; Dr. S. Anderson and Mr. W. Sha for critical reading of the manuscript; Dr. E. Sadler and Ms. L. Westfield for the oligonucleotides; and Ms. S. Hahn for secretarial assistance.

Received for publication 10 July 1989.

References

1. Kronenberg, M., G. Siu, L. E. Hood, and N. Shastri. 1986. The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *Annu. Rev. Immunol.* 4:529.
2. Chou, H. S., S. J. Anderson, M. C. Louie, S. A. Godambe, M. R. Pozzi, M. A. Behlke, K. Huppi, and D. Y. Loh. 1987. Tandem linkage and unusual RNA splicing of the T-cell receptor β-chain variable-region genes. *Proc. Natl. Acad. Sci. USA.* 84:1992.
3. Wade, T., J. Bill, P. C. Marrack, E. Palmer, and J. W. Kappler. 1988. Molecular basis for the nonexpression of Vβ17 in some strains of mice. *J. Immunol.* 141:2165.
4. Behlke, M. A., D. G. Spinella, H. S. Chou, W. Sha, D. L. Hartl, and D. Y. Loh. 1985. T-cell receptor β-chain expression: Dependence on relatively few variable region genes. *Science (Wash. DC).* 229:566.
5. Behlke, M. A., H. S. Chou, K. Huppi, and D. Y. Loh. 1986. Murine T-cell receptor mutants with deletions of β-chain variable region genes. *Proc. Natl. Acad. Sci. USA.* 83:767.
6. Chou, H. S., C. A. Nelson, S. A. Godambe, D. D. Chaplin, and D. Y. Loh. 1987. Germline organization of the murine T cell receptor β-chain genes. *Science (Wash. DC).* 238:545.
7. Haqqi, T. M., S. Banerjee, W. L. Jones, G. Anderson, M. A. Behlke, D. Y. Loh, H. S. Luthra, and C. S. David. 1989. Identification of T-cell receptor Vβ deletion mutant mouse strain AU/SsJ (H-2^q) which is resistant to collagen-induced arthritis. *Immunogenetics.* 29:180.
8. Kappler, J. W., N. Roehm, and P. Marrack. 1987. T cell tolerance by clonal elimination in the thymus. *Cell.* 49:273.
9. Marrack, P., and J. Kappler. 1988. T cells can distinguish between allogeneic major histocompatibility complex products on different cell types. *Nature (Lond.).* 332:840.
10. Kappler, J. W., U. Staerz, J. White, and P. C. Marrack. 1988. Self-tolerance eliminates T cells specific for Mls-modified products of the major histocompatibility complex. *Nature (Lond.).* 332:35.
11. MacDonald, H. R., R. Schneider, R. K. Lees, R. C. Howe, H. Acha-Orbea, H. Festenstein, R. M. Zinkernagel, and H. Hengartner. 1988. T-cell receptor Vβ use predicts reactivity and tolerance to Mls^a-encoded antigens. *Nature (Lond.).* 332:40.
12. Pullen, A. M., P. Marrack, and J. W. Kappler. 1988. The T-cell repertoire is heavily influenced by tolerance to polymorphic self-antigens. *Nature (Lond.).* 335:796.
13. Abe, R., M. S. Vacchio, B. Fox, and R. J. Hodes. 1988. Preferential expression of the T-cell receptor Vβ3 gene by Mls^c reactive T cells. *Nature (Lond.).* 335:827.
14. White, J., A. Herman, A. M. Pullen, R. Kubo, J. W. Kappler, and P. Marrack. 1989. The Vβ-specific superantigen staphylococcal enterotoxin B: stimulation of mature T cells and clonal deletion in neonatal mice. *Cell.* 56:27.

15. Kappler, J. W., T. Wade, J. White, E. Kushnir, M. Blackman, J. Bill, N. Roehm, and P. Marrack. 1987. A T cell receptor V_β segment that imparts reactivity to a class II major histocompatibility complex product. *Cell.* 49:263.
16. Singer, P. A., R. J. McEvilly, D. J. Noonan, F. J. Dixon, and A. N. Theofilopoulos. 1986. Clonal diversity and T-cell receptor β-chain variable gene expression in enlarged lymph nodes of MRL-*lpr/lpr* lupus mice. *Proc. Natl. Acad. Sci. USA.* 83:7018.
17. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
18. Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA.* 74:560.
19. Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294.
20. Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12:7035.
21. Anderson, S. J., H. S. Chou, and D. Y. Loh. 1988. A conserved sequence in the T-cell receptor β-chain promoter region. *Proc. Natl. Acad. Sci. USA.* 85:3551.
22. Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6.
23. Patten, P., T. Yokota, J. Rothbard, Y. Chien, K. Arai, and M. M. Davis. 1984. Structure, expression and divergence of T-cell receptor β-chain variable regions. *Nature (Lond.).* 312:40.
24. Blankenstein, T., F. Bonhomme, and U. Krawinkel. 1987. Evolution of pseudogenes in the immunoglobulin V_H-gene family of the mouse. *Immunogenetics.* 26:237.
25. Haqqi, T. M., S. Banerjee, G. D. Anderson, and C. S. David. 1989. RIII S/J (H-2^r): an inbred mouse strain with a massive deletion of T cell receptor V_β genes. *J. Exp. Med.* 169:1903.