Identification and Physical Mapping of A Polymorphic Human T Cell Receptor Vβ Gene with a Frequent Null Allele

By Patrick Charmley, Kai Wang, Leroy Hood, and Deborah A. Nickerson

From the Division of Biology, California Institute of Technology, Pasadena, California 91125

Summary

Germline variation in genes that encode the human T cell receptors (TCRs) may have an important influence in shaping the immune T cell repertoire. In this report we describe a frequent null allele of the human Vβ18 gene, resulting from a nucleotide substitution that creates a stop codon (CGA↔TGA). Approximately 11% of the population tested was homozygous for this null allele, indicating that this is a frequent "hole in the repertoire." We confirmed that there is a greatly reduced (undetectable) level of Vβ18 mRNA in peripheral blood lymphocytes from an individual homozygous for this null allele. In addition, all heterozygous individuals expressed detectable levels of only the functional Vβ18 allele in their peripheral blood lymphocytes. Two other DNA polymorphisms were identified in Vβ18, one of which would result in an amino acid substitution in an expressed Vβ18 gene. Genotypes for all three of these Vβ18 DNA polymorphisms were determined in a group of unrelated individuals. Statistical analyses of the associations between alleles of the Vβ18 polymorphisms and those of other DNA polymorphisms in the TCR β locus suggested a close physical proximity between the Vβ18 gene and the 3′ end of the Cβ2 region. This localization of human Vβ18 had been previously predicted by the sequence homology between human Vβ18 and mouse Vβ14, a V gene segment previously mapped to 3′ of the mouse Cβ genes. We confirmed this localization of the human Vβ18 gene by isolating a cosmid clone that contains both the Vβ18 and Cβ2 gene segments. Mapping by restriction enzyme digestion and by the polymerase chain reaction indicated that the Vβ18 gene segment is approximately 9 kb 3′ of the Cβ2 gene, making this the only known human Vβ gene 3′ of the Cβ region.

A diverse repertoire of TCRs is generated by a variety of somatic mechanisms, including the random rearrangement of gene segments, N region diversification, and junctional diversity. In addition to these somatic effects, allelic variations arising from DNA polymorphisms in the TCR genes also contribute to germline diversity in the population (1). In the past, RFLPs detected with probes from TCR V genes were used as indirect measures of DNA polymorphism within the V genes (2–6). More recently, however, direct sequence analysis has demonstrated the presence of germline DNA polymorphisms in the TCR that result in amino acid substitutions. Specifically, Robinson (7) has reported a polymorphism in the Vβ1 gene segment that results in a change from glutamine to histidine, while Li et al. (8) reported two polymorphisms in the Vβ6.7 gene that result in a serine to arginine and also a glycine to glutamic acid substitution. Although the ultimate impact of these or other allelic variations on the TCR repertoire remains to be determined, the analysis of TCR polymorphisms may provide a greater understanding of the potential relationship between DNA polymorphisms in the TCR and susceptibilities to certain autoimmune diseases (reviewed in references 1 and 9).

To further assess the extent of TCR germline polymorphism, we have compared the sequence of TCR V gene segments from multiple individuals. In this study, we report a striking example of a DNA polymorphism in the Vβ18 gene that results in the introduction of a termination codon (the Vβ18 gene has also been referred to as the Vβ20 gene in another nomenclature by Kimura et al. [9a]). 11% of the individuals analyzed were homozygous for this null allele, and therefore could not express a functional Vβ18 protein product. Analysis of Vβ18 expression using RNA from PBL confirmed this relationship between Vβ18 RNA message and the stop codon. Additionally, we have also obtained the entire genomic sequence of the Vβ18 gene, and have determined its physical location to be at the 3′ end of the C region genes.

 J. Exp. Med. © The Rockefeller University Press • 0022-1007/93/01/0135/09 $2.00
Volume 177 January 1993 135–143

This localization is opposite of other known human Vβ genes, but is syntenic with the location of a homologous TCR V gene segment (Vβ14) previously mapped in the mouse.

Materials and Methods

Human DNA and RNA. A panel of DNAs isolated from the lymphoblastoid cell lines of six unrelated Caucasians was used for polymorphism screening, and was kindly provided by R. A. Gatti. The DNA of 76 Caucasian parents and certain families from Centre d'Etude du Polymorphisme Humaine (CEPH) (10), pedigrees were used for polymorphism typing. For the Vβ18 expression analysis, RNA was isolated from PHA-stimulated PBL from several small families (kindly provided by R. Spielman, University of Pennsylvania, Philadelphia, PA) using the acid guanidinium thiocyanate-phenol-chloroform extraction method (11).

Oligonucleotides. Amplification primers and ligation probes were synthesized using a DNA synthesizer (384 or 391; Applied Biosystems, Inc., Foster City, CA). The sequences of amplification primers and ligation probes used in polymorphism typing of the Vβ18 gene segment are shown in Table 1. The 5′ allele-specific ligation probes were biotinylated as previously described (12), and the adjoining 3′ reporter probes were phosphorylated using "5′-phosphate-on" according to the manufacturer's instructions (Clonetech, Palo Alto, CA). Ligation probes were purified by reverse-phase HPLC, and the phosphorylated 3′ reporter probes were enzymatically labeled with digoxigenin as previously described (13).

DNA Amplification. DNA samples were amplified by the PCR. For determining the Vβ18 genotypes by restriction enzyme digestion, PCR was performed using a DNA thermal cycler (Perkin-Elmer Cetus, Emeryville, CA) in a reaction mix containing reaction buffer (Amersham Corp., Arlington Heights, IL), nucleotides (final concentration, 20 μM each), DNA polymerase (Hot Tub; Amersham Corp.) (final concentration, 10 U/ml), 0.3 μM each of primer Vβ18-3 and Vβ18-4 (Table 1), and 10–100 ng of genomic DNA. The amplification profile was 35–40 cycles of 1 min at 94°C, 1 min at 60°C, and 2 min at 72°C.

For determining the Vβ18 genotypes by OLA, DNA samples were amplified in a 96-well microtiter plate thermal cycler (MJ Research, Watertown, MA). Genomic DNA (10 ng) was mixed with a standard PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, and 0.001% gelatin) containing 40 μM of each of the four deoxynucleotide triphosphates, 0.5 μM of each primer, and 10 U/ml of Taq polymerase (Perkin-Elmer Cetus) (14). The amplification profile was 40 cycles of 20 s at 93°C, 40 s at 55°C, and 90 s at 72°C.

For PCR-based mapping of the H29 cosmid, 50 ng of cosmid DNA was mixed with the same mixture of PCR buffer, nucleotides, and enzyme as described directly above for amplifying DNA in a 96-well format. The primer pair used for amplification was 5′-CTGTTTCCCTGAAGATTGAG-3′ (3′ end of Cβ2) and 5′-ACTGGGTGGAGGATATTTGTAAG-3′ (3′ end of Vβ18). This mix was then subjected to 40 cycles of 1 min at 93°C, 1 min at 55°C, and 2 min at 72°C in a DNA thermal cycler (Perkin-Elmer Cetus). The amount of time at 72°C was extended by 5 s after each cycle and is essential to amplifying larger DNA products.

DNA Sequencing. For purposes of DNA sequencing, a large-scale PCR (200 μl) was performed according to the conditions described above. To eliminate an additional manipulation, mineral oil was not used during the amplification so that ethanol precipitation was done in the same tube as used for amplification. The concentrated PCR products were purified by electrophoresis through 1% low-melting point agarose (Bethesda Research Laboratories, Gaithersburg, MD), and the amplified products excised from the gel. 10 μl of the PCR product in the gel slice was used directly as the template for sequencing in 96-well assay plates using the Sequenase sequencing kit (U.S. Biochemical Corp., Cleveland, OH) (15). The sequencing primers used were the same as those used for the original amplification. Sequence from the Vβ18 gene region was obtained using the H29 cosmid. Sequence-specific oligonucleotide primers were used to sequence across both strands of the cosmid template (~2 μg of cosmid DNA template per labeling reaction). Sequencing was done using the Sequenase sequencing kit, and the annealing and radiolabeling steps were performed at 37°C (15).

Restriction Digestion Analysis. Restriction digestion of the PCR products was performed without further purification of the DNA. 10 μl of the amplified DNA was digested with 10 U of either KpnI or Sau3A (Bethesda Research Laboratories). Only the addition of BSA (final concentration, 100 μg/ml) was necessary for complete digestion with KpnI, whereas addition of the manufacturer's 10× digestion buffer was required for digestion with Sau3A. Restriction enzyme–digested PCR products were subjected to electrophoresis in a 3% agarose gel (including 2% Nusieve agarose; FMC Bioproducts, Rockland, ME) containing ethidium bromide, and visualized by UV illumination.

Oligonucleotide Ligation Assay (OLA).[1] Analyses of amplified DNA samples using primers Vβ18-1 and -2 for each of the Vβ18 polymorphisms were performed using an OLA detection system previously described in detail (13, 14). In this assay, each sample was analyzed by a ligation reaction for each allele. Each reaction contained one of the 5′ biotinylated probes specific for that allelic form of the polymorphism, the common 3′ reporter probe labeled with digoxigenin, an aliquot of the amplified DNA target, and T4 DNA ligase. If the probes were hybridized to a perfectly matched target, the biotinylated allele specific and reporter probes were joined by the ligase. To determine whether the ligation reaction had been successful, the reaction was transferred to a streptavidin plate to capture the biotinylated probe. The plate was washed and then assayed for the presence of the covalently attached reporter molecule using an enzyme-linked immunoassay system (13).

Cosmid Identification and Restriction Digest Mapping. A human cosmid library constructed in a modified pWE15 vector (16) and containing 15 genome equivalents was screened for Vβ18. The cosmid H29 was identified by hybridization (final stringency was 0.5× SSC at 65°C) to a PCR-amplified Vβ18 gene segment that had been cloned into the M13 phage. The H29 cosmid was mapped with EcoRI, HindIII, BamHI, and XbaI using a standard protocol of partial digestion, and then sequential hybridization with vector-specific oligonucleotides (final stringency wash of 6× SSC at 65°C), and to TCR Cβ and Vβ18 probes.

Vβ18 RNA Expression. For PCR analysis of RNA, cDNA was synthesized from 1–2 μg of total RNA by random hexamer priming using a cDNA synthesis kit (Amersham Corp.). Approximately 25 ng of the cDNA was PCR amplified with 25 cycles using a 5′ Vβ18 and 3′ Cβ together with a 5′ and 3′ Cα primer pair, as listed in Fig. 6, *b* and *c*, legends. The products were subjected to electrophoresis at 150 V through 2% agarose and vacuum blotted with 0.4 N NaOH onto Genetran nylon membrane (Plasco, Marlboro, MA). The membrane was hybridized overnight to a Cβ and Cα DNA probe that were ^{32}P radiolabeled separately by random priming. These probes had been combined as a 50:50 mixture based on the amount of incorporated ^{32}P. The hybridized blot was then

[1] *Abbreviation used in this paper:* OLA, oligonucleotide ligation assay.

washed several times up to a final stringency of 0.1× SSC at 65°C. The PCR products were detected by autoradiography.

Linkage Disequilibrium Analysis. Statistical analysis of linkage disequilibrium between loci was tested using the Q statistic (17). This statistic measures the total linkage disequilibrium between the alleles at two loci by summating the differences between the observed frequencies of the allele combinations and those frequencies expected if one assumed random association of alleles at the two loci. Linkage disequilibrium was also assessed using a normalized measure for disequilibrium, D′ (18), which describes the extent of disequilibrium in terms of the fraction of maximum (±1.0) linkage disequilibrium possible for that locus pair. Since genetic phase was not established, individuals who were heterozygous at both loci in a pair-wise comparison were not included in the linkage disequilibrium calculation.

Results

Genomic Sequence of the Vβ18 Region. To obtain the genomic sequence of the human Vβ18 region, a cosmid (H29) was isolated by hybridization to a Vβ18 probe. Using a series of primer-directed sequencing walks along both strands of the cosmid DNA, the entire Vβ18 coding sequence was obtained as listed in Fig. 1. This sequence extends from the leader region, through the first intron and the Vβ18 gene segment, and ends within the 3′ noncoding sequence. The 3′ noncoding sequence also includes the recombination signal (heptamer-nonamer sequences) for somatic DNA rearrangement. One of the interesting features of the human Vβ18 gene sequence is its similarity to mouse Vβ14 (19). This similarity across species is 77% in the coding regions, and 63% overall.

Identification of Sequence Polymorphisms in the Human Vβ18 Gene. Direct sequencing of the amplified germline Vβ18 gene segment from multiple persons revealed three DNA polymorphisms in the Vβ18 gene segment. The positions of the DNA polymorphisms in the Vβ18 gene segment are shown in Fig. 1. All three polymorphisms were single nucleotide substitutions involving C to T base transitions. From the amino acid translations, two polymorphisms, one at nucleotide 524 and the other at nucleotide 620, would have significance with regard to the Vβ18 protein product. Most importantly, the polymorphism at position 524 leads to the introduction of a stop codon in the Vβ18 gene sequence. The nucleotide substitution at position 620 would result in a proline to serine amino acid substitution in the gene product.

Genotyping of the Vβ18 Polymorphisms. Genotypes for the

```
    MetLeuCysSerLeuLerAlaLeuLeuLeuGlyThrPhePhe
A.  ATGCTCTGCTCTCTCCTTGCCCTTCTCCTGGGCACTTTCTTTGgtgggtgccctcctctctgcaagcctcttttctcatc   80
B.  --------------------------------------------------------------------------------
C.  --------------------------------------------------------------------------------

A.  ctcacctccaccccaatcccgtgaatgcagagtctgagctgttggtttgacactcttgagtctattcctgattgtctgga  160
B.  --------------------------------------------------------------------------------
C.  --------------------------------------------------------------------------------

A.  acagatatgggaatttagccttaggggctgatctgatgattcttagttctctctatctcctgagaaggagttcgttctgt  240
B.  --------------------------------------------------------------------------------
C.  --------------------------------------------------------------------------------

A.  tggtcttccctggcccaggatgctttgcatccttggctcccccagtctactccccatgatggattccaggtgcctgtgtc  320
B.  --------------------------------------------------------------------------------
C.  --------------------------------------------------------------------------------

A.  atgcaagagcttgagctgcccctctgtctggggaaggcaagccaggctggggtgctcctctgtaaatggcactttatctt  400
B.  --------------------------------------------------------------------------------
C.  --------------------------------------------------------------------------------

           GlyValArgSerGlnThrIleHisGlnTrpProAlaThrLeuValGlnProValGlySerProLeuSerLeuGlu
A.  tccacagGGGTCAGATCTCAGACTATTCATCAATGGCCAGCGACCCTGGTGCAGCCTGTGGGCAGCCCGCTCTCTCTGGA  480
B.  --------------------------------------------------------------------------------
C.  --------------------------------------------------------------------------------

     CysThrValGluGlyThrSerAsnProAsnLeuTyrTrpTyrArgGlnAlaAlaGlyArgGlyLeuGlnLeuLeuPheTyr
A.  GTGCACTGTGGAGGGAACATCAAACCCCAACCTATACTGGTACCGACAGGCTGCAGGCAGGGGCCTCCAGCTGCTCTTCT  560
B.  --------------------------------------------------------------------------------
C.  -------------------------------------------T------------------------------------
                                               Stop

      SerValGlyIleGlyGlnIleSerSerGluValProGlnAsnLeuSerAlaSerArgProGlnAspArgGlnPheIle
A.  ACTCCGTTGGTATTGGCCAGATCAGCTCTGAGGTGCCCCAGAATCTCTCAGCCTCCAGACCCCAGGACCGGCAGTTCATC  640
B.  --------------------------------------------------------------------------------
C.  -----------------------------------------------------------T--------------------
                                                               Ser

    LeuSerSerLysLysLeuLeuLeuSerAspSerGlyPheTyrLeuCysAlaTrpSer
A.  CTGAGTTCTAAGAAGCTCCTCCTCAGTGACTCTGGCTTCTATCTCTGTGCCTGGAGTgtcacactgagctgggtggggca  720
B.  --------------------T-----------------------------------------------------------
C.  --------------------T-----------------------------------------------------------

A.  gacatctgtgcaaaaaccccaccctctcctgagccctaaccatactccccagggccttcacttagggactgggtggagga  800
B.  -----------------------------------
C.  -----------------------------------

    tatttgtaagtaggttctgactgctggccaatgtgaagc                                           839
```

Figure 1. Complete genomic DNA and amino acid sequence of the human Vβ18 gene segment. The DNA and amino acid sequence polymorphisms are shown at positions 524, 620, and 661. The actual DNA sequence of the three Vβ18 haplotypes observed is indicated by the *A*, *B*, and *C* lines (-, identical sequence). The heptamer-nonamer somatic recombination signals found in the 3′ noncoding region are italicized and bolded. PCR primers used to amplify this gene segment are underlined. The double-underlined DNA sequences were found to be different compared with the original reported cDNA sequence of HUT102 (20), and were not found to represent additional polymorphism. These sequence data are available from EMBL/GenBank/DDBJ under accession number Z13967 HSTCRVB18.

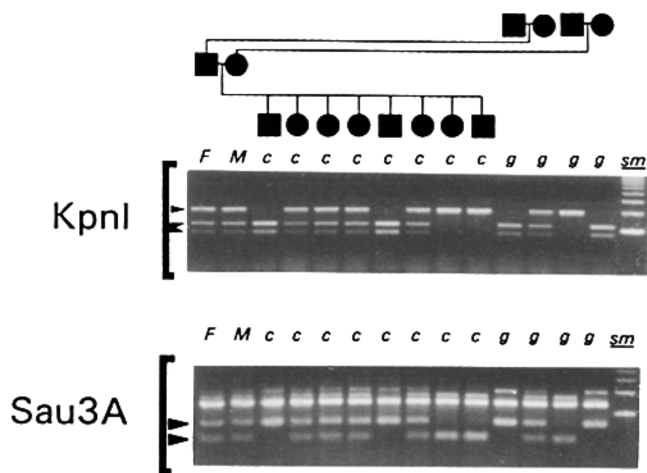

Figure 2. Mendelian segregation of the KpnI and Sau3A Vβ18 DNA polymorphisms. The Vβ18 gene was amplified using primers Vβ18-3 and -4 from the individuals in a three-generation pedigree (*F*, father; *M*, mother; *c*, child; *g*, grandparent). The PCR products were digested with either the KpnI or Sau3A restriction enzyme and genotypes were determined based on the polymorphic DNA fragments observed. The pointers on the left indicate the allelic restriction fragments, with the joined pointers (after KpnI digestion) indicating a single allele (though the lower fragment is noticeably faint in the heterozygotes). Sau3A digestion also resulted in a monomorphic DNA fragment present in all individuals. On the far right, *sm* indicates the 123-bp ladder used as a size marker.

Vβ18 polymorphisms were determined for a larger group of individuals. First, according to the DNA sequences surrounding the polymorphisms at nucleotide positions 524 and 620 (Fig. 1), these polymorphisms would result in altering restriction enzyme recognition sites. The base change at position 524 alters a KpnI restriction site, and the DNA polymorphism at position 620 changes a Sau3A restriction site, as demonstrated in Fig. 2. Genotypes for all three Vβ18 polymorphisms were obtained using the semi-automated OLA typing system. Analysis of the inheritance patterns for these three DNA polymorphisms revealed proper Mendelian segregation, as shown in Fig. 3. The segregation of the Vβ18 genotypes was completely compatible with cosegregation (i.e., no recombinants) with haplotypes of the TCR β chain previously determined in these individuals (6).

Within our tested population (152 chromosomes), the nucleotide changes at positions 524 and 620 had identical allele frequencies of 68% for the C substitution and 32% for the T substitution. 11% of the individuals tested were homozygous for the T substitution at nucleotide 524, and therefore, based on the presence of the stop codon, these 11% of the individuals would not produce a functional Vβ18 protein product. The allele frequency for the silent base substitution at nucleotide 661 was slightly different from those found for the other two polymorphisms (524 and 620), i.e., at position 661 the allele frequency for the C substitution was 53% and the T substitution was 47%. Even though one of the substitutions, 524, leads to a stop codon, the genotype distributions for all three polymorphisms were consistent with Hardy-Weinberg equilibrium.

Linkage Disequilibrium Analysis. We statistically examined the extent of association between alleles of the Vβ18 polymorphism(s) and alleles of other previously reported TCR β DNA polymorphisms (Table 2). Because of complete linkage disequilibrium among the three point mutations within the Vβ18 gene (D′ = 1.00), only three DNA haplotypes were observed (as shown in Fig. 1). The extent of linkage disequilibrium was also measured between the Vβ18 polymorphisms and other biallelic polymorphisms previously studied in this same population of CEPH parents (6, 14). Although the alleles of the Vβ18 polymorphisms did not show statistical association with other Vβ gene polymorphisms (Table 2), polymorphisms at the 3′ end of the TCR β gene complex (Cβ2 region) were significantly nonrandomly associated with the 524 and 620 Vβ18 polymorphism(s) (p <0.0001).

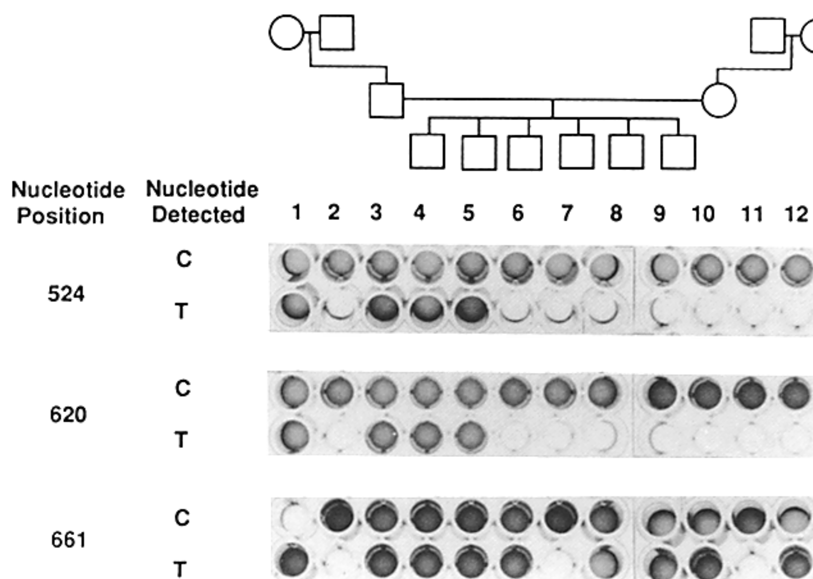

Figure 3. A composite figure assembled from photographs of microtiter plates containing OLA genotypes that show Mendelian segregation of the three Vβ18 sequence polymorphisms. Microtiter wells containing the reporter group digoxigenin form a colored product (*dark wells*) and indicate a perfect nucleotide match between the ligating probes and the amplified DNA target.

Table 1. *DNA Sequences for the Amplification Primers and Ligation Probes Used in Typing the Vβ18 DNA Polymorphisms*

Amplification primers	Sequences	
Vβ18-1	GAGTGCACTGTGGAGGGAA	
Vβ18-2	TCACATTGGCCAGCAGTCAG	
Vβ18-3	ATTCATCAATGGCCAGCGAC	
Vβ18-4	GGAGCTTCTTAGAACTCAG	
Ligation probes*		
Vβ18-524	B-CCCCAACCTATACTGGTACC	pGACAGGCTGCAGGCAGGG-D
	B-CCCCAACCTATACTGGTACT	
Vβ18-620	B-GAATCTCTCAGCCTCCAGAC	pCCCAGGACCGGCAGTTCAT-D
	B-GAATCTCTCAGCCTCCAGAT	
Vβ18-661	B-TGAGTTCTAAGAAGCTCCTC	pCTCAGTGACTCTGGCTTCTAT-D
	B-TGAGTTCTAAGAAGCTCCTT	

* B, biotin; p, phosphorylated; D, digoxigenin.

Table 2. *Analysis of Linkage Disequilibrium between Vβ18 Polymorphisms and Other Polymorphisms in the TCR β Gene Complex*

Vβ18-524/620* and:	Haplotypes	D'‡	Q (χ^2)§	Vβ18-661 and:	Haplotypes	D'	Q (χ^2)
Vβ1	140	0.04	0.1	Vβ1	130	0.07	0.1
Vβ8	116	0.18	1.3	Vβ8	92	−0.20	2.9
Vβ11	114	−0.23	1.9	Vβ11	94	−0.20	2.5
Cβ2-1	118	−0.28	2.2	Cβ2-1	102	−0.30	4.5
Cβ2-5	102	−1.0	31.4	Cβ2-5	94	−0.35	8.7
Cβ2-6	108	0.69	24.8	Cβ2-6	98	0.27	6.4
Cβ2-7	98	−1.0	27.2	Cβ2-7	92	−0.37	9.0
Vβ18-661	84	1.0	31.2	Vβ18-524/620*	84	1.0	31.2

The order of the polymorphisms compared in this table (top to bottom) is the same as the physical order of the polymorphisms along the chromosome (centromere to telomere) (6, 14).

* The Vβ18 position 524 and 620 polymorphisms are synonymous for this analysis since there was a singular correspondence between alleles at these loci (haplotypes, 86; D', 1.0; Q, 86.0).

‡ For D', a value of 1.00 (or −1.00) indicates maximum linkage disequilibrium; a value of 0.0 would indicate complete random association between the loci.

§ For the Q statistic (χ^2 distribution, 1-degree freedom), Q >15 corresponds to p <0.0001, which rejects the null hypothesis of random association between the alleles at that pair of loci.

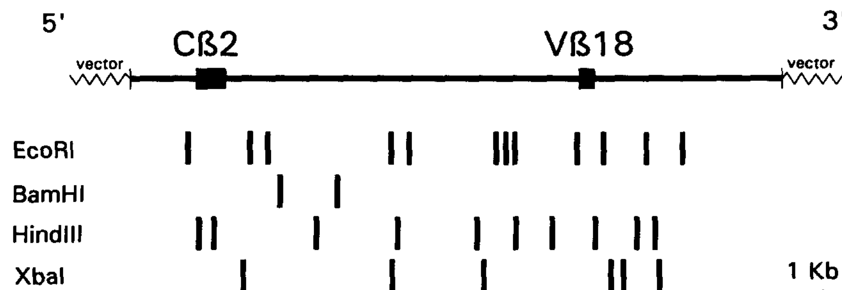

Figure 4. Restriction map of the cosmid insert containing Vβ18 and Cβ2 genes. Cosmid H29 was shown by hybridization to contain both the Vβ18 and Cβ gene segments. A restriction map of the cosmid was constructed by partial digestion and hybridization with either the gene segments or vector sequences. Restriction sites for EcoRI, BamHI, HindIII, and XbaI are shown as vertical lines below the cosmid.

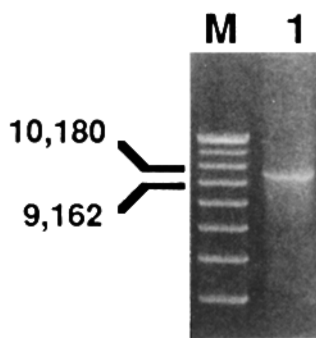

Figure 5. Ethidium bromide–stained gel containing the amplified product from the H29 cosmid using a primer located at the 3′ end of the Cβ2 gene and one located 3′ of the Vβ18 gene. Lane *M* contains a size marker consisting of a 1-kb ladder. The sizes of the two size marker bands flanking the amplified Cβ2-Vβ18 product are indicated to the left.

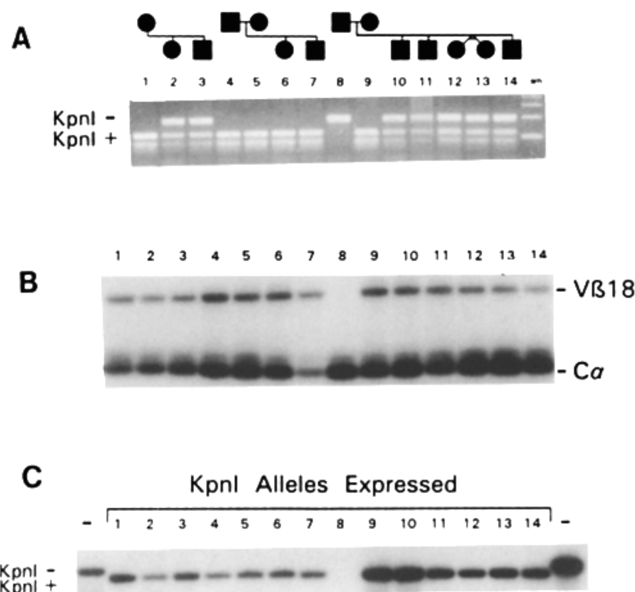

Figure 6. Vβ18 DNA genotypes and RNA expression in PBL of 14 individuals. (*A*) PCR-amplified Vβ18 gene segment from genomic DNA was digested with KpnI. The DNA polymorphism that results in the stop codon also destroys this KpnI site. The family structures are shown above. (*B*) The somatically rearranged Vβ18 gene was amplified from the PBL RNA (cDNA) of these same individuals. The forward primer Vβ18-3 and a reverse Cβ primer (5′-AGTCATTGAGGGCGGGCTG-3′) produced the larger Vβ18-containing product, and each tube also contained a TCR Cα primer pair (5′-CAGCATTATTCCAGAAGACAC-3′, and 5′-CCTCAGCTGGACCACAGC-3′) as a positive control for amplification. (*C*) The rearranged Vβ18 gene was amplified from cDNA by PCR using the forward primer Vβ18-3 and a reverse Cβ primer (5′-GGCAGACAGGACCCCTTG-3′); the Cα primer pair listed above was also included. The PCR products were digested with KpnI restriction enzyme to confirm that the Vβ18 null allele (no KpnI site) was not expressed. As a control for the Vβ18-Cβ KpnI$^-$ allele, the outer lanes (–) contain the PCR product to which KpnI enzyme was not added. The Cα PCR-positive controls in each lane are not shown.

Physical Mapping of the Vβ18 Cosmid. Based on the linkage disequilibrium results (Table 2), we sought to confirm the suggested 3′ location of the Vβ18 gene relative to Cβ2. The H29 cosmid that contained Vβ18 was tested for hybridization to 22 other Vβ families and Cβ. Only hybridization to a Cβ probe was detected. PCR was used to confirm that this cosmid contained both the Vβ18 and Cβ genes. Physical mapping of the H29 cosmid with four restriction enzymes and hybridization with Vβ18 and Cβ probes indicated that the Vβ18 gene was located within 15 kb of the Cβ2 gene (Fig. 4). Previous studies have shown that the mouse Vβ14 gene (gene counterpart of human Vβ18) has an opposite transcriptional orientation to that of other Vβ genes as well as the D, J, and C gene segments (20). By using PCR with the H29 cosmid, we have determined the transcriptional orientation of human Vβ18 relative to the Cβ2 gene segment. DNA amplification with a primer located just 3′ of the end of the Cβ2 gene, in combination with a primer located just downstream of the Vβ18 gene, indicated that the Vβ18 gene segment is located ∼9 kb 3′ of the Cβ2 gene (Fig. 5). Since these PCR primers must face each other to amplify the intervening DNA sequence, these data also indicate that the transcriptional orientation for human Vβ18 is opposite to that of the Cβ2 gene, and therefore similar to the orientation of the mouse Vβ14 and Cβ2.

It is worth noting that several different pairs of amplification primers from the 3′ end of the Cβ2 gene and the 3′ end of the Vβ18 gene could be used to obtain a similarly sized PCR product from cosmid DNA (data not shown). In addition, we were also able to sequence the ends of the 9.5-kb PCR product (Fig. 5) in low-melt agarose (see Materials and Methods), which verified the origins of this PCR product, i.e. Cβ2 at the 5′ end and Vβ18 at the 3′ end (data not shown).

Vβ18 RNA Expression in Peripheral Blood. As a result of the null allele for Vβ18 in the population, we sought to confirm the importance of this polymorphism in the expression level of Vβ18 in individuals homozygous for the Vβ18 stop codon. Several members from three families were tested by KpnI restriction enzyme analysis of genomic DNA for the Vβ18 null allele (KpnI$^-$). Representatives of each of the three genotypes were identified (Fig. 6 *A*) and PCR was used to amplify the Vβ18 RNA (cDNA) from all these individuals. Based on the intensity of the Vβ18 PCR product, and the control Cα product in each lane, Vβ18-Cβ message was detected in the peripheral blood of all individuals, except the person homozygous for the stop codon (Fig. 6 *B*). As even stronger evidence of the relationship between the stop codon and Vβ18 expression, the amplified Vβ18-Cβ message was digested with KpnI (Fig. 6 *C*). The results show that in the eight individuals studied who possessed the null allele (either heterozygous or homozygous individuals), only the Vβ18 KpnI$^+$ allele was expressed at detectable levels in the peripheral blood.

Discussion

Allelic variation in the germline DNA of the TCR genes has been suspected to influence the expressed TCR repertoire by altering amino acids that change the TCR's ability to recognize antigen and/or MHC. DNA polymorphisms in TCR genes could change the expressed T cell repertoire either qualitatively or quantitatively through the process of clonal selection and expansion in the thymus or periphery. It is also possible that DNA polymorphisms could influence the quan-

tity of an expressed TCR protein by altering regulatory elements, or through translational changes such as creating a stop codon. The potential for functional relevance of TCR polymorphisms has been the basis for many disease association studies that have compared the frequency of various TCR DNA polymorphisms in autoimmune and control populations (reviewed in references 1 and 9).

In this study, we describe an extreme example of a functionally relevant DNA polymorphism that results in the presence of a stop codon in the Vβ18 gene segment. This polymorphism was present in 54% of the population that we studied. Importantly, 11% of the individuals tested were homozygous for this null allele, indicating a lack of cell surface expression of the Vβ18 TCR protein. Coincidentally, there were two other DNA polymorphisms in the Vβ18 gene, one of which would have resulted in an amino acid change from proline to serine, in the absence of the stop codon.

The observation of complete allelic association between the Vβ18 polymorphisms could have implications in terms of the origins of these polymorphisms. Since the Vβ18 alleles at positions 524 and 620 were correlated completely, this could be due to the simultaneous "creation" of the polymorphisms in a single ancestral genetic event. A more speculative but immunological explanation for the complete allelic correlation is evolutionary selection at the protein level since the alleles of both polymorphisms would affect expression and possibly structure. At present time, whenever the serine residue is coded for at the second polymorphic site, the termination codon is encoded at the first polymorphic site. The serine/proline substitution site would occur within the fourth hypervariable region in the TCR β chain (21) and might be involved in antigen/MHC contact. Based on the location of this amino acid substitution, and the radical nature of the substitution (22), generous speculation might suggest that the Vβ18 haplotype with the stop codon and serine residue has been selected to prevent protein expression of the Vβ18 with a serine residue at this position.

The strong disequilibrium between the Cβ-KpnI site (Cβ2-5; Table 2) and the Vβ18 polymorphisms has relevance to disease association studies since previous studies that have used the Cβ-KpnI RFLP have therefore indirectly measured the Vβ18 polymorphism association with systemic lupus erythematosus, multiple sclerosis, and myasthenia gravis (23, 24). Also, the lack of any detectable disequilibrium between the Vβ18 polymorphisms and the Cβ-BglII RFLP (Cβ2-1), as well as the other more distant 5′ Vβ RFLPs (Table 2), implies that the many positive disease associations found with those RFLPs (1, 9) cannot be directly explained by the Vβ18 stop codon. Vice versa, studies reporting no disease association with the previously reported Vβ and Cβ-BglII RFLPs in Table 2 (except Cβ-KpnI) cannot conclude that the Vβ18 polymorphisms are not involved in the disease susceptibility.

Physical mapping of the human and mouse TCR β gene complexes has shown a remarkable conservation in the organization of homologous Vβ genes between these two species (19, 25). The human Vβ18 gene is most similar to the mouse Vβ14 gene based on the high nucleic acid homology between these gene segments (77%). Therefore, since the mouse Vβ14 has been shown to be the only V gene located 3′ of the TCR β C region (26), it was predicted that the human Vβ18 gene would be located 3′ of the C region (19). This was suggested by linkage disequilibrium (Table 2) and confirmed by isolating a cosmid that contained both the human Vβ18 and C region Cβ2 gene. Restriction enzyme and PCR mapping of this cosmid showed Vβ18 to be ∼9 kb 3′ of the Cβ2 gene segment. This 9-kb distance in humans is similar to the 10-kb distance between Cβ2 and Vβ14 in mice (26). Multiple PCR analyses also showed that the human Vβ18 has an opposite transcriptional polarity relative to the TCR β D, J, and C region genes and other Vβ genes, as shown for mouse Vβ14 (26). Thus, it may similarly be expected that human Vβ18 undergoes a chromosomal inversion to juxtapose Vβ18 with the D, J, and C region genes before successful transcription.

Since the human Vβ18 gene was originally cloned from an acute lymphocytic leukemia cell line (20), and since it is 3′ of Cβ as well as being polymorphic for a stop codon, it may be questioned as to the quantitative contribution of a functional Vβ18 protein product to the total TCR β repertoire. Choi et al. (27) estimated that Vβ18 (Vβ20 in their nomenclature) accounted for ∼3.4% of the total Vβ expression after anti-CD3 stimulation of peripheral T cells from one individual, while another group failed to detect Vβ18 expression in peripheral blood from one donor (28). Vβ18 is a single member subfamily (29), so there are no other Vβ18 subfamily members that might subsume the role of Vβ18 and prevent this hole in the TCR repertoire caused by the null allele.

Our result showing greatly reduced (undetectable) Vβ18 RNA expression in the peripheral blood of the null allele in all eight individuals who possessed the null allele (Fig. 6, *c* and *b*) is consistent with the mechanisms of T cell selection that occur in the thymus. The positive or negative selection of T cells in the thymus is dependent on the specific α and β proteins on the cell's surface. T cells that rearrange and transcribe the nonfunctional Vβ18 gene segment could not be positively selected because they would be unable to assemble a TCR α/β protein heterodimer. Thus, these Vβ18-rearranged cells could not be expected to populate the peripheral blood and immune system to an appreciable level. In an individual homozygous for the Vβ18 null allele, a very low level of rearranged Vβ18 message in PBL could be accounted for by the infrequent additional rearrangement that can occur on the other chromosome 7 homologue in T cells that express some other Vβ gene on their cell surface. The failure to detect Vβ18-Cβ RNA message from the null allele is also consistent with what has been previously seen in mice with a similar Vβ stop codon (30, and see below [BALB/c Vβ19]).

While other polymorphic stop codons have not yet been reported in human TCR genes, there are two examples of polymorphic stop codons in mouse TCR Vβ genes. Mouse Vβ19 is expressed in the SJL strain, but the BALB/c strain has deleted a single base in the leader exon that results in a frameshift mutation (30). A second example of a polymorphic stop codon in a mouse TCR β gene is mouse Vβ17. Certain

strains of mice possess a TCR β haplotype that encodes a Vβ17 gene with a single mutation in the coding region that results in a stop codon (31). Other strains of mice with a functional Vβ17 gene have their Vβ17-bearing T cells clonally eliminated in those mice that express a certain self-superantigen and the MHC class II molecule IE (32). Wade et al. (31) speculated that although self-tolerance mechanisms would seem to be able to control such autoimmune cells reactive to the self-superantigen, perhaps evolutionary pressures have acted to provide an additional means of eliminating such autoreactive T cells from a portion of the population, via a stop codon. It is therefore interesting to note that the T cells bearing human Vβ18 appear to be stimulated by the *Staphylococcus aureus* enterotoxin superantigens (27). This parallel between these mouse and human Vβ genes allows similar speculation on whether past selective pressures on humans could have had an importance in the prevalence of the polymorphic stop codon in the human Vβ18 gene.

We thank Ms. Wendy Ankener and Ms. Patsy Byers for technical assistance and oligonucleotide syntheses, Drs. Jim Brawley, Patrick Concannon, Bill Funkhouser, Andy Herman, Ben Koop, Eric Milner, Jerry Nepom, and Dave Ostrov for helpful discussions and manuscript review.

This work was supported by National Institutes of Health grants HG-00084 and DK-41347, the Whittier Foundation, and National Science Foundation grant DIR 8809710.

Address correspondence to P. Charmley, Virginia Mason Research Center, 1000 Seneca Street, Seattle, WA 98101. K. Wang's, L. Hood's, and D. A. Nickerson's present address is the Department of Molecular Biotechnology (GJ-10), University of Washington School of Medicine, 4909 25th Ave. N.E., Seattle, WA 98195.

Received for publication 23 June 1992 and in revised form 2 October 1992.

Note added in proof: Malhotra et al. (33) have recently reported the detection, by a different approach and technology, of the mutation at nucleotide position 524 in the Vβ18 gene (the Vβ20 gene according to their nomenclature).

References

1. Posnett, D.N. 1990. Allelic variations of human TCR V gene products. *Immunol. Today.* 11:368.
2. Concannon, P., R.A. Gatti, and L.E. Hood. 1987. Human T cell receptor Vβ gene polymorphism. *J. Exp. Med.* 165:1130.
3. Ball, E.J., L. Dombrausky, M. Hoover, J.D. Capra, and P. Stastny. 1987. Restriction fragment length polymorphism of the human T cell receptor α gene. *Immunogenetics.* 26:48.
4. Grier, A.H., M.P. Mitchell, and M.A. Robinson. 1990. Polymorphism in human T cell receptor alpha chain variable region genes. *Exp. Clin. Immunogenet.* 7:34.
5. Chan, A., R.-P. Du, M. Reis, L.M. Meske, M. Sheehy, E. Baillie, and T.W. Mak. 1990. Human T cell receptor Vα gene polymorphism. *Exp. Clin. Immunogenet.* 7:26.
6. Charmley, P., A. Chao, P. Concannon, L. Hood, and R.A. Gatti. 1990. Haplotyping the human T-cell receptor β-chain gene complex by use of restriction fragment length polymorphisms. *Proc. Nat. Acad. Sci. USA.* 87:4823.
7. Robinson, M.A. 1989. Allelic sequence variations in the hypervariable region of a T-cell receptor β chain: Correlation with restriction fragment length polymorphism in human families and populations. *Proc. Natl. Acad. Sci. USA.* 86:9422.
8. Li, Y., P. Szabo, M.A. Robinson, B. Dong, and D.N. Posnett. 1990. Allelic variations in the human T cell receptor Vβ6.7 gene products. *J. Exp. Med.* 171:221.
9. Concannon, P. 1992. T-cell receptor gene polymorphisms: applications in the study of autoimmune disease. *In* Manual of Clinical Immunology. American Society for Microbiology, Washington, D.C. 885–889.

9a. Kimura, N., B. Toyonaga, Y. Yoshikai, R.-P. Du, and T.W. Mak. 1987. Sequences and repertoire of the human T cell receptor α and β chain variable region genes in thymocytes. *Eur. J. Immunol.* 17:375.

10. Dausset, J., H. Cann, D. Cohen, M. Lathrop, J.-M. Lalouel, and R. White. 1990. Centre d'etude du polymorphisme humain (CEPH): collaborative genetic mapping of the human genome. *Genomics.* 6:575.
11. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156.
12. Landegren, U., R. Kaiser, J. Sanders, and L. Hood. 1988. A ligase-mediated gene detection technique. *Science (Wash. DC).* 241:1077.
13. Nickerson, D.A., R. Kaiser, S. Lappin, J. Stewart, L. Hood, and U. Landegren. 1990. Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. *Proc. Natl. Acad. Sci. USA.* 87:8923.
14. Nickerson, D.A., C. Whitehurst, C. Boysen, P. Charmley, R. Kaiser, and L. Hood. 1992. Identification of clusters of biallelic polymorphic sequence-tagged sites (pSTSs) that generate highly informative and automatable markers for genetic linage mapping. *Genomics.* 12:377.
15. Kretz, K.A., S.C. Geoffrey, and S.O. O'Brien. 1989. Direct sequencing from low-melt agarose with Sequenase. *Nucleic Acids Res.* 17:5864.
16. Lai, E., Wang, K., Avdalovic, N., and L. Hood. 1991. Rapid restriction map constructions using a modified pWE15 cosmid

vector and a robotic workstation. *Biotechniques.* 11:212.

17. Hedrick, P.W., G. Thompson, and W. Klitz. 1986. Evolutionary genetics: HLA as an exemplary system. *In* Evolutionary Processes and Theory. S. Karlin, and E. Nevo, editors. Academic Press, New York. 583–606.
18. Lewontin, R.C. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics.* 49:49.
19. Wilson, R.K., E. Lai, P. Concannon, R.K. Barth, and L.E. Hood. 1988. Structure, organization, and polymorphism of murine and human T-cell receptor α and β chain gene families. *Immunol. Rev.* 101:149.
20. Leiden, J.M., and J.L. Strominger. 1986. Generation of diversity of the β chain of the human T-lymphocyte receptor for antigen. *Proc. Natl. Acad. Sci. USA.* 83:4456.
21. Jores, R., P.M. Alzari, and T. Meo. 1990. Resolution of hypervariable regions in T-cell receptor β chains by a modified Wu-Kabat index of amino acid diversity. *Proc. Natl. Acad. Sci. USA.* 87:9138.
22. Bohinski, R.C. 1983. Modern Concepts in Biochemistry, 4th ed. Allyn and Bacon, Inc., Boston.
23. Perl, A., J.P. Divincenzo, P. Gergely, J.J. Condemi, and G.N. Abraham. 1989. Detection and mapping of polymorphic *KpnI* alleles in the human T-cell receptor constant beta-2 locus. *Immunology.* 67:135.
24. Oksenberg, J.R., C. Newby Gaiser, L.L. Cavalli-Sforza, and L. Steinman. 1988. Polymorphic markers of human T-cell receptor alpha and beta genes. Family studies and comparison of frequencies in healthy individuals and patients with multiple sclerosis and myasthenia gravis. *Hum. Immunol.* 22:111.
25. Lai, E., P. Concannon, and L. Hood. 1988. Conserved organization of the human and murine T-cell receptor β-gene families. *Nature (Lond.).* 331:543.
26. Malissen, M., C. McCoy, D. Blanc, J. Trucy, C. Devaux, A.M. Schmitt-Verhulst, F. Fitch, L. Hood, and B. Malissen. 1986. Direct evidence for chromosomal inversion during T-cell receptor β-gene rearrangements. *Nature (Lond.).* 319:28.
27. Choi, Y., B. Kotzin, L. Herron, J. Callahan, P. Marrack, and J. Kappler. 1989. Interaction of *Staphylococcus aureus* toxin "superantigens" with human T cells. *Proc. Natl. Acad. Sci. USA.* 86:8941.
28. Moss, P.A.H., R.J. Moots, W.M.C. Rosenberg, S.J. Rowland-Jones, H.C. Bodmer, A.J. McMichael, and J.I. Bell. 1991. Extensive conservation of α and β chains of the human T-cell antigen receptor recognizing HLA-A2 and influenza A matrix peptide. *Proc. Natl. Acad. Sci. USA.* 88:8987.
29. Robinson, M.A. 1991. The human T cell receptor β-chain gene complex contains at least 57 variable gene segments. Identification of six Vβ genes in four new gene families. *J. Immunol.* 146:4392.
30. Louie, M.C., C.A. Nelson, and D.Y. Loh. 1989. Identification and characterization of new murine T cell receptor β chain variable region (Vβ) genes. *J. Exp. Med.* 170:1987.
31. 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.
32. Kappler, J.W., T. Wade, J. White, E. Kushnir, M. Blackman, J. Bill, N. Roehm, and P. Marrack. 1987. A T cell receptor Vβ segment that imparts reactivity to a class II major histocompatibility complex product. *Cell.* 49:263.
33. Malhotra, U., R. Spielman, and P. Concannon. 1992. Variability in T cell receptor Vβ gene usage in human peripheral blood lymphocytes. Studies of identical twins, siblings, and insulin-dependent diabetes mellitus patients. *J. Immunol.* 149:1802.