The EMBO Journal vol.8 no.12 pp.3741 - 3748, 1989

The smaller human V_H gene families display remarkably little polymorphism

Iñaki Sanz[1], Patrick Kelly, Carol Williams, Stephen Scholl, Philip Tucker and J.Donald Capra

Department of Microbiology, The University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA

[1]Present address: The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78284, USA

Communicated by T.H.Rabbitts

We report the nucleotide sequence of 30 distinct human V_H gene segments from the V_HIV, V_HV and V_HVI gene families. When these sequences were compared to previously published sequences from these smaller human V_H families a surprisingly low level of polymorphism was noted. Two V_HIV gene segments from unrelated individuals were identical to two previously published V_HIV sequences. Five V_HV sequences were identical and seven V_HVI gene segments were identical. Where differences were found between the sequences, allele specific oligonucleotide probes were used to verify the germline nature of the change and to test for segregation in several large kindreds. These data provide evidence that at least some human V_H gene segments are remarkably stable.

Key words:

Introduction

An antibody heavy chain variable region (V_H) is encoded by three discrete gene segments named variable (V_H), diversity (D) and joining (J_H). In the germline configuration each of these genes is separated from the others by several kb of intervening DNA. In the antibody-producing B cell, these gene segments are juxtaposed in a well defined two step recombinational process producing a functional V_H gene (Tonegawa, 1983).

V_H gene segments constitute a unique genetic system. It has been estimated that up to 1000 V_H genes are present in the murine genome (Livant *et al.*, 1986). The murine V_H genes can be divided by amino acid and nucleotide sequence homology into at least 11 V_H families (Brodeur and Riblet, 1984). Arbitrarily those genes sharing at least 80% sequence identity are considered members of the same family. With some exceptions murine V_H genes are highly polymorphic, and this extensive polymorphism is well documented by both restriction fragment length polymorphism (RFLP) and nucleotide sequence analysis (Rathbun *et al.*, 1987). Two general types of variation have been described in the murine V_H system. In one, a V_H gene segment in one strain has no obvious counterpart or allele in a second strain. Thus, idiotypic or RFLP variation can be ascribed to the presence or absence of a locus. In the second type, clear allelism is obvious and the differences between strains reflect a few nucleotides. Many inherited idiotypic systems in the murine inbred strains can be related to such variations among the V_H gene segments. For example, the NP^a/NP^b (Bothwell *et al.*, 1981) and arsonate (Siekevitz *et al.*, 1983) idiotypes are well known examples of such V_H gene segment polymorphisms.

The human V_H locus spans ~2 megabases on chromosome 14 (14q32), with ~200 V_H genes which can be grouped into six discrete families (Kodaira *et al.*, 1986; Schroeder *et al.*, 1987; Berman *et al.*, 1988; Humphries *et al.*, 1988). Interestingly, RFLP as well as amino acid and nucleotide sequence comparison of V_H genes from unrelated individuals suggest that the level of polymorphism of some human V_H genes, even if significant in the context of non-polymorphic genes, may be remarkably low as compared with many studies of inbred strains of mice (Sanz *et al.*, 1988; Van Dijk *et al.*, 1989). Indeed, there is a growing list of expressed autoantibodies that have been shown to have identical nucleotide sequences even when derived from unrelated individuals suggesting that the same germline gene may exist in a large proportion of the species (Dersimonian *et al.*, 1987; Chen *et al.*, 1988; Pascual *et al.*, 1989; Sanz *et al.*, 1989).

The human V_HI, V_HII and V_HIII families were described several years ago by sequencing monoclonal paraproteins from lymphoid malignancies. V_HI and III are the largest and most heterogeneous families with 50 - 100 members (Matthyssens and Rabbitts, 1980. Rechavi *et al.*, 1982). V_HII contains perhaps 20 genes. On average, these families contain up to 30 - 40% pseudogenes. In the past four years, three new families have been defined. V_HIV was first recognized by Baer *et al.* 1985, in their analysis of an inversion of human chromosome 14. Lee *et al.* (1987)

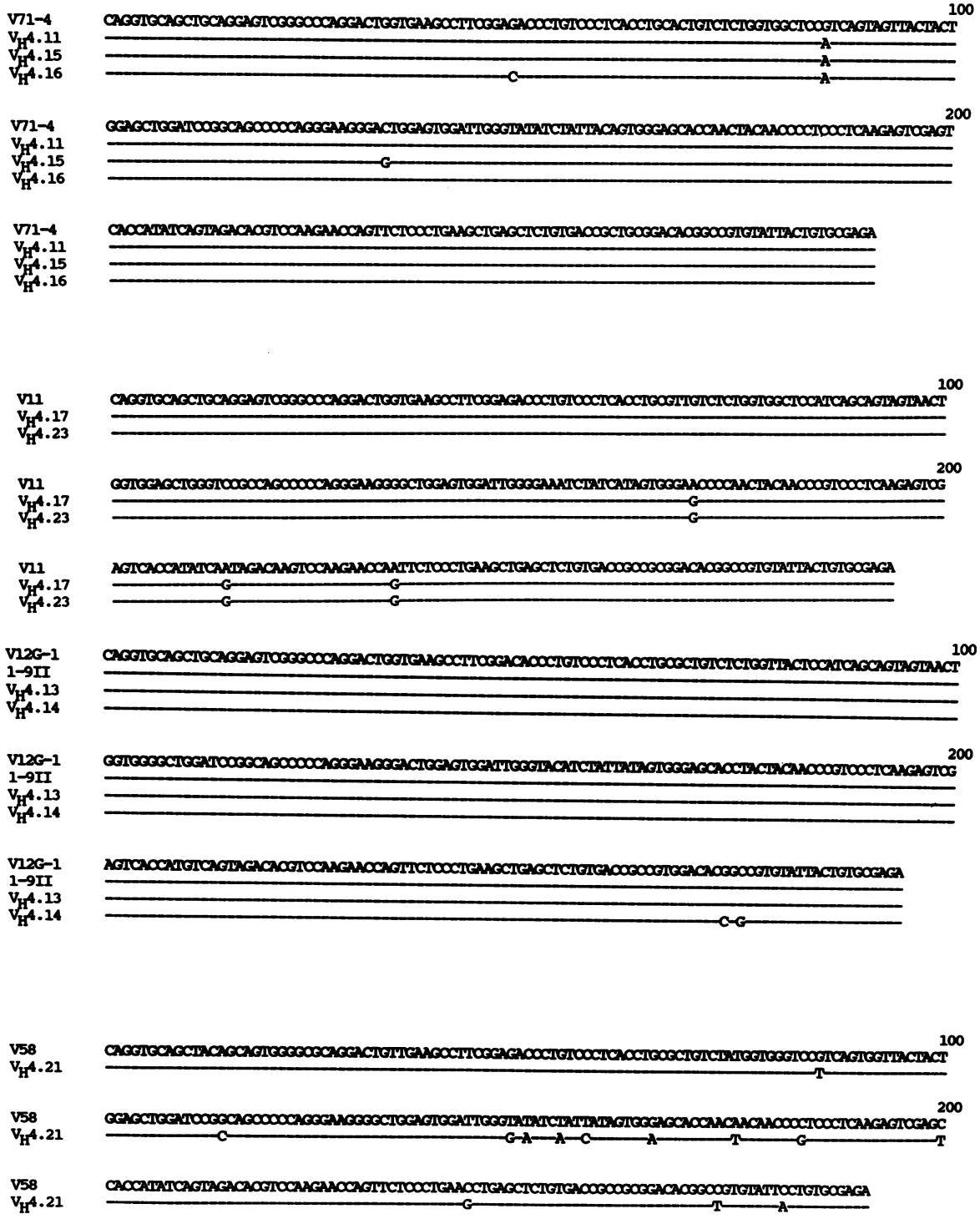
Table I. Comparison of V_HIV genes sequenced in this study to those of Lee *et al.* (1987)

	V71-4	V11	V12G-1	V58	V2-1	V79	V71-2
V_H4.11	99.6*	93.1	94.5	94.1	94.8	93.8	99.3
V_H4.12	97.6	93.5	94.9	92.7	95.9	94.2	98.0*
V_H4.13	93.5	93.5	[100.0]*	90.3	93.5	94.2	94.2
V_H4.14	92.8	92.8	99.3*	89.6	92.8	93.5	93.5
V_H4.15	99.3*	93.1	94.5	94.1	94.8	93.8	99.3
V_H4.16	99.3*	92.7	94.8	93.8	94.5	93.4	99.0
V_H4.17	93.1	99.0*	94.5	90.7	93.5	98.3	93.8
V_H4.18	94.5	92.8	93.8	91.0	99.7*	93.5	94.9
V_H4.19	93.1	97.6	94.5	90.7	93.5	[100.0]*	93.8
V_H4.21	92.4	91.7	90.0	95.5*	91.7	92.4	92.0
V_H4.22	94.8	92.8	95.2	92.0	96.6*	93.5	95.2
V_H4.23	93.1	99.0*	94.5	90.7	93.5	98.3	93.8

Comparisons shown in Figure 1 are marked with an asterisk in the Table. These represent the closest homology of the V_HIV genes sequenced in this study to the V_HIV genes sequenced by Lee *et al.* (1987).

subsequently cloned and sequenced the germline V_HIV genes. This is a very homogeneous family with internal homology being at least 91.5% when the most distant member of the family (V58) are compared to either a consensus sequence or to another V_HIV gene (Lee *et al.*, 1987), V_HV has been thought to be composed of only a single functional gene and two pseudogenes (Humphries *et al.*, 1988) and V_HVI has but a single member that has been mapped as the D−J_H proximal V_H gene ~90 kb from the D locus (Berman *et al.*, 1988; Buluwela and Rabbitts, 1988; Schroeder *et al.*, 1988). It has been difficult, however, to assess variations in specific human V_H genes as the 'classical' V_H gene families (V_HI, V_HII and V_HIII) are large and direct comparisons between individuals are difficult. The recent description of three 'smaller' V_H gene families (<10 members), however, makes such comparisons practical.

A thorough knowledge of the human V_H complex seems necessary to help answer several fundamental questions. (i) What genetic mechanisms are responsible for the evolution of the V_H complex in an outbred population? (ii) Are germline sequences preserved because of an evolutionary advantage to recognize self antigens, external antigens or because of neutral evolution? Here we explore these issues and present data concerning the nucleotide sequence of 30



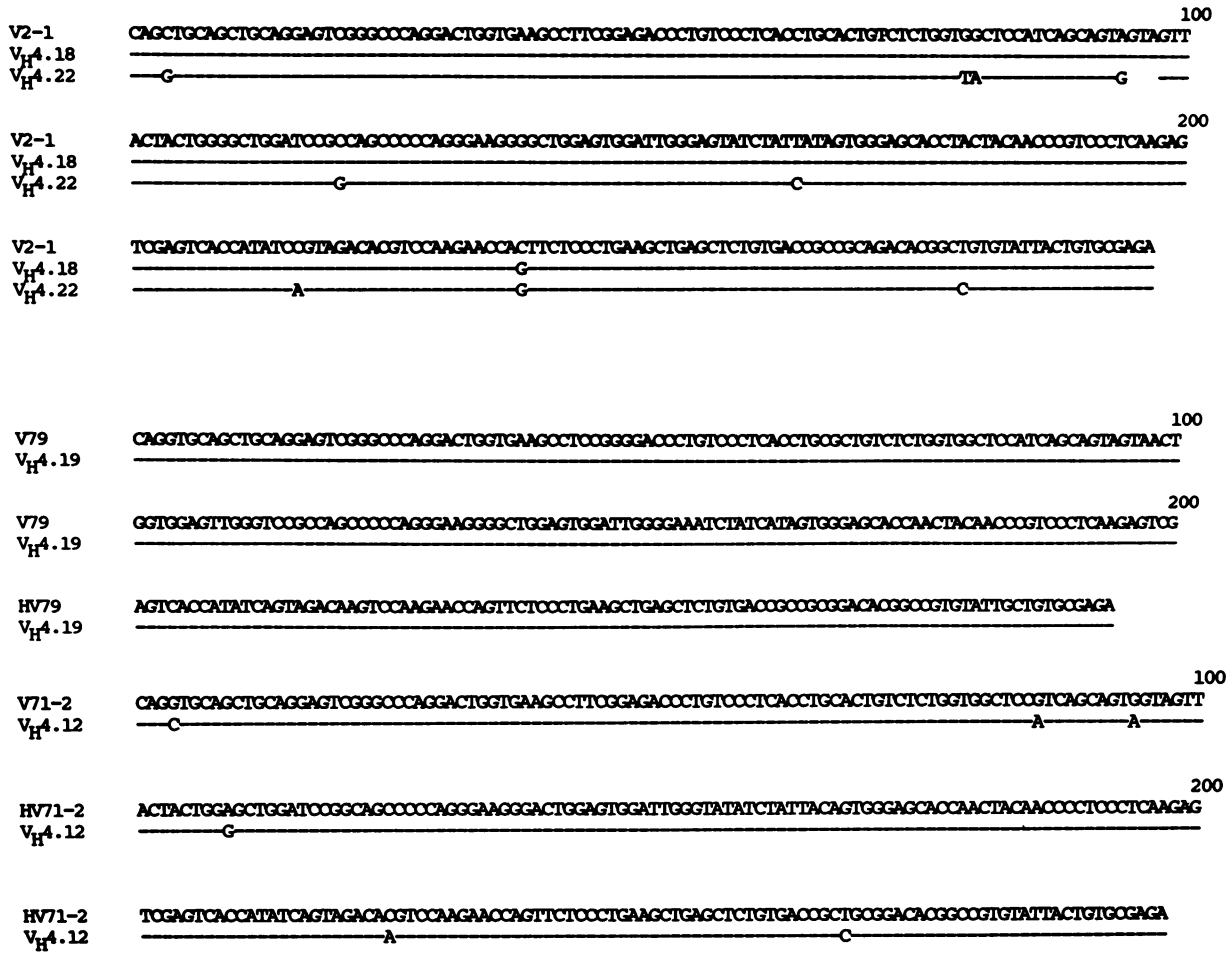


Fig. 1. The nucleotide sequences of 12 human V_HIV family germline genes isolated from two unrelated Caucasian donors. The sequences are compared with the V_HIV family members described by Lee *et al.* (1987) and one by Berman *et al.* (1-9II) (Berman *et al.*, 1988). Only differences from Lee *et al.* are shown. In each comparison we chose the closest V_H gene segment described by Lee *et al.* (1987). Note that V_H4.21 is probably not allelic with V58 but V58 is the closest published V_HIV sequence to V_H4-21.

human V_H genes of the 'smaller' human V_H families V_HIV, V_HV and V_HVI. We document that these smaller V_H families display remarkably little polymorphism.

Results

The nucleotide sequence of a total of 12 V_HIV genes from two genetically unrelated individuals reveals some identities between the two individuals and identities to previous sequenced V_HIV gene segments

Amplification experiments with the V_HIV specific oligonucleotide primers were performed with genomic DNA from two unrelated healthy Caucasian blood donors. These individuals will be designated hereafter as V_H4.1 and V_H4.2. Ten V_HIV positive clones from the V_H4.1 donor and five from the V_H4.2 donor were randomly selected and sequenced (Table I and Figure 1). From donor V_H4.1, three clones (V_H4.11A, V_H4.11B and V_H4.11C) were identical to each other and will be discussed further as V_H4.11. Two others (V_H4.16A and V_H4.16B) had but a single difference with V_H4.11 (99.7% identity) and were identical to each other and will be further referred to as V_H4.16. Another, V_H4.15, had a single difference with both V_H4.11 and V_H4.16. Two other clones (V_H4.13 and V_H4.14) were 99.3% identical to each other. These findings could be due to multiple copies of the same gene from the same PCR reaction, with an error introduced by the *in vitro* amplification procedure. Alternatively, this individual might be heterozygous for these genes such that, e.g. V_H4.16 is the allelic counterpart of V_H4.11. These possibilities were explored with allele specific oligonucleotide probes (see below).

The V_HIV family has been estimated to contain at least nine members. The sequence of seven of them have been reported (Lee *et al.*, 1987, Berman *et al.*, 1988). When the V_HIV genes sequenced in this study were compared to the V_HIV genes previously reported, it was possible to identify putative allelic counterparts for all but V58 (Table I). Three of our sequences were 99% or more identical to V71-4, (V_H4.11, V_H4.15 and V_H4.16, marked with an asterisk in the V71-4 column of Table I). The three identical clones V_H4.11A, B and C were 99.6% identical to V71-4; V_H4.15, V_H4.16A and V_H4.16B were 99.3% identical to V71-4 (see Figure 1). V_H4.17 (donor V_H4.1) was 99% identical to V11 and 100% identical with V_H4.23 from donor V_H4.2. In Table I and Figure 1, V_H4.11 and V_H4.23 are listed separately although they are identical because they were isolated from different individuals. In all other cases (i.e. V_H4.11A, B and C; V_H4.16A and B) only one sequence is presented as these are likely duplicate clones. V_H4.13 was 100% identical to V12-G1 (Lee *et al.*, 1987) and 1-9II (Berman *et al.*, 1988). V_H4.14 was 99.3%

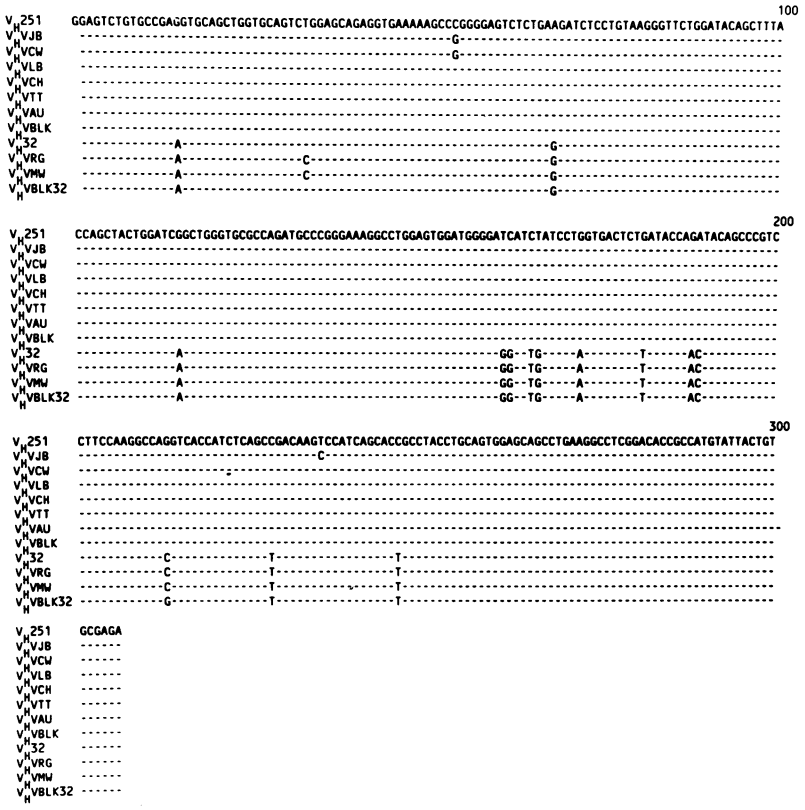


Fig. 2. The nucleotide sequence of 10 V_HV genes isolated from 10 Caucasian donors are compared to V_H 251 (Humphries *et al.*, 1988). V_H32, another V_HV member is also included. Only differences from V_H251 are shown. V_HBLK and V_HBLK32 are from the same donor.

identical to these same two V_HIV genes. $V_H4.18$ was 99.7% identical to V2-1 and $V_H4.19$ was 100% identical to V79. V71-2 is the only reported V_HIV gene with a six nucleotide insertion in the first hypervariable region (Lee *et al.*, 1987). The same insertion is also present in $V_H4.12$ but not in any of the other genes. $V_H4.12$ is 98% identical to V71-2.

Five V_HIV positive clones were sequenced (Table I) from donor $V_H4.2$ but only three unique sequences were identified. $V_H4.21A$ and B were identical to each other (and will be further referred to as $V_H4.21$) as well as $V_H4.22A$ and B (and will be referred to as $V_H4.22$). As previously mentioned, $V_H4.23$ was 100% identical to $V_H4.17$ and therefore 99.3% identical to V11. The best homology found for $V_H4.22$ was 96.6% with V2.1 and 95.5% for $V_H4.21$ with V58.

Not counting identical sequences from the same donor, 12 germline V_H genes from two unrelated Caucasians were sequenced and compared to the genes reported by Lee *et al.* (1987) and Berman *et al.* (1988). Out of 3179 nucleotides analyzed, there were 37 variations for a total conservation of 98.8% throughout the V_HIV family. The $V_H4.21$ gene from the second donor probably constitutes a V_H58 subfamily not found thus far in other individuals studied. When only the $V_H4.1$ donor is considered, only 14 variations were found in 2312 nucleotides analyzed representing only 0.6% divergence in eight sequences.

Out of the 37 mutations found, 17 were transitions and 20 were transversions. Eleven mutations were found in hypervariable regions as opposed to 26 in the framework regions. The ratio of replacement to silent mutations was 2.6 in the hypervariable regions versus 1.6 in the framework regions. It should be noted that all of the genes sequenced contain an open reading frame (ORF) suggesting that the frequency of pseudogenes in the V_HIV family might be extremely low. This constitutes a significant departure from the estimated number of 40% pseudogenes found in other human V_H families (Kodaira *et al.*, 1986).

The nucleotide sequences of 11 V_HV genes show remarkable identities

The V_HV gene family has been recently described (Shen *et al.*, 1987; Humphries *et al.*, 1988; Makar *et al.*, 1988). It consists of three members: V_H251, V_H32 and V_H15 of which only V_H251 is reported to contain an ORF. V_H32 and V_H15 contain termination codons and, therefore, are considered pseudogenes. Using specific primers for the leader intron and 3′ recombination sequences, we amplified two (V_H251 and V_H32) of the three V_HV genes and cloned and sequenced the allelic counterparts of V_H251 from five Caucasians, one Oriental and an African Black and the allelic counterpart of V_H32 from two Caucasians and one African Black (Figure 2). Striking identities were found among these unrelated individuals. Six genes (V_HLB, V_HCH, V_HTT, V_HAU and V_HBLK) were 100% identical to each other but different in two positions from V_H251. V_HJB has two additional nucleotide differences from V_H251. V_HCW shared with V_HJB the silent substitution. Similarly, the V_H32 genes were remarkably similar. These sequences from three unrelated persons from different racial backgrounds were 99.6% identical to V_H32 with only one substitution each. Furthermore, V_HRG and V_HMW were 100% identical to each other. These three sequences maintain a cytosine residue at the end of framework three that had

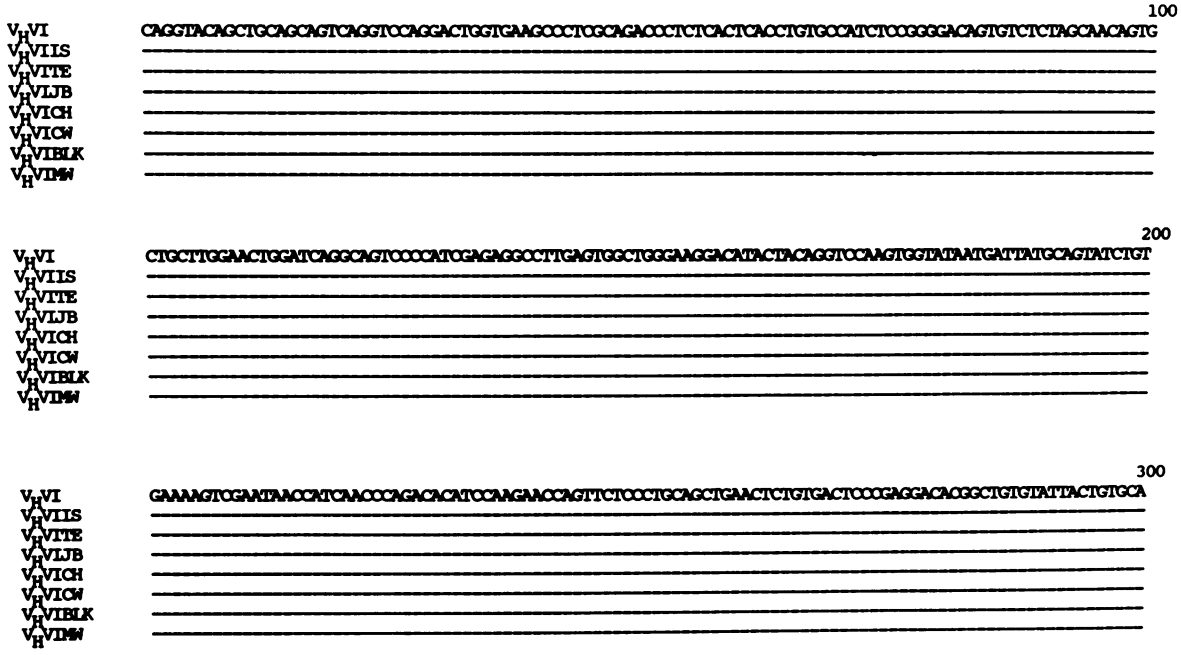


Fig. 3. The nucleotide sequence of seven V_HVI germline genes isolated from seven unrelated donors. The sequences are compared with the V_HVI germline genes reported by others (Schroeder *et al.*, 1987; Berman *et al.*, 1988; Buluwela and Rabbitts, 1988).

been reported to be deleted in V_H32 but is present in V_H251 and V_H15.

These and other considerations led us to re-examine the original V_H251 and V_H32 clones and we determined that there were two errors in the reported V_H251 sequence and one in V_H32. As such, V_H32 does not contain a termination codon and is a functional gene. Indeed, we have recently determined that an antibody to the acetylcholine receptor isolated from an EBV-transformed B cell line from a patient with autoimmune myasthenia gravis utilizes an unmutated V_H32 gene (V.Pascual; A.Lefvert and J.D.Capra, unpublished observations). The sequences shown in Figure 2 reflect corrections of V_H251 and V_H32.

Seven V_HVI gene segments from unrelated individuals are identical to each other and to three previously sequenced V_HVI gene segments

The V_HVI gene has recently been shown to be composed of a unique member that is the most D−J_H proximal human V_H gene segment (Berman *et al.*, 1988; Buluwela and Rabbitts, 1988; Schroeder *et al.*, 1988). The physical distance between the V_HVI gene and the J_H locus is thought to be less than 90 kb. In this study of human V_HVI polymorphisms, we have sequenced the V_HVI germline genes from seven unrelated individuals and found that in all cases, they are 100% identical to one another as well as to the three V_HVI nucleotide sequences reported in the literature (Figure 3). In each case five separate clones were independently sequenced and no variations were noted. While it is possible that in each individual the same allele was amplified, cloned and sequenced, these data argue that the V_HVI gene segment, at least in these seven individuals is not polymorphic.

Allele specific oligonucleotide analysis

The finding that essentially the entire V_HIV, V and VI families can be amplifed with unique oligonucleotides constructed to the leader introns and heptamer/nonamer

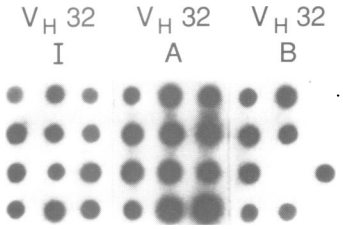


Fig. 4. Allele specific oligonucleotide dot blots detect polymorphism in the V_H32 genes. The V_HV family was amplified by PCR as described. Probe 'I' is an internal oligonucleotide which serves as an amplification control (it was also used in colony screening to detect genes to be sequenced). V_H32A and V_H32B represent allele specific oligoncleotide probes (ASOP) which distinguish the T/C difference in V_H32 at position 32. The ASOP also contains an 'A' at position three in the oligonucleotide (position 15 in Figure 2) as otherwise it would detect V_H251 genes. As seen in this example, 12 members of this panel were positive for the A allele ('T' at position 32) but only 8 out of 12 were positive for the B allele ('C' at position 32). Subsequent studies showed this variation to segregate in families.

introns provided the opportunity to assess each of the sequence variations noted in order to determine (i) if these represented sequence errors introduced either during the amplification or sequencing procedure, or (ii) represented point mutations in the individual donating the DNA, and/or (iii) represented polymorphisms in the human V_H complex. Accordingly, allele specific oligonucleotides were made to seven regions of the V_HIV gene that were different either between the genes sequenced in our own laboratory or between these genes and those sequenced elsewhere. Additionally, four different pairs of allele specific oligonucleotides (eight probes) were constructed to distinguish differences in the V_HV family and two (four probes) within the V_HVI families (these latter two were constructed to differences noted between V_HVI genes in humans and chimpanzees; K.Meek and J.D.Capra, unpublished.

The results from one typical analysis is illustrated in Figure

Table II. Allele specific oligonucleotide analysis of V_HIV/V_HV genes

V_H family	Locus	Position	Distinction		Our population[c]	
			A versus B		A versus B	
IV	V71-4	85	G	A	0	12
IV	V71-4	48	G	C	7	12
IV	11	170	A	G	0	12
IV	11	214	A	G	0	12
IV	V71-2	267	T	C	4	12
V[a]	251	53	C	G	5	17[b]
V	251	235	T	C	2	17[b]
V	32	32	T	C	10	17[b]

[a]In addition, previous errors in V_H251 and V_H32 were tested (prior to reassessing the sequences) and were not detected in 12 normals.
[b]These differences were shown to segregate in large kindreds.
[c]These data are expressed as individuals, not alleles.

4 and a more extensive list is shown in Table II. An internal oligonucleotide controls for amplification in each instance. Certain oligonucleotide probes are positive in some members of a family but not in others. All told these probes were used in 12 normal unrelated individuals and the V_HV probes were further studied in five extended families. Almost all of the variations noted in the sequences reported in this paper represent true polymorphisms (i.e. C/G V_H251 position 53, T/C V_H32 position 32, T/C V71.2 position 267) whereas some (i.e. V71.4 G/A position 85) are either point mutations in the individual, polymorphisms not present in our panel or represent sequencing errors. Overall, however, the extent of polymorphism was remarkably small. None of the variations noted between human and chimpanzee V_HVI genes were detected in 16 normals of various racial and ethnic backgrounds. Interestingly, none of the differences between the sequences reported here and the V_HIV sequences reported by Lee *et al.* (1987) could be shown to be present in the germline of our population. Yet the variations between our sequences were found in the 12 (in some cases 17) unrelated individuals in our panel. These results argue that the differences between our sequences are not technical but rather represent polymorphic differences within the population. Some of the nucleotide differences between our sequences and those of Lee *et al.* (1987) could represent racial differences as our panel was largely Caucasian.

Discussion

An understanding of the extent of polymorphism of V_H genes will be important in order to appreciate the evolutionary role of these loci as well as the genetic mechanisms which shape them. Furthermore, an understanding of the extent of polymorphism may shed light as to the possible association of V_H genes with at least some autoimmune diseases. This issue has been studied extensively in inbred strains of mice but has been difficult to address in the human. It is generally assumed that V_H genes are highly polymorphic in rodents as determined by RFLP studies as well as nucleotide sequence analysis. RFLP analysis, even if not as sensitive as sequencing, is able to determine major differences for most V_H gene families in inbred strains. More detailed and sensitive studies have been performed in at least three separate murine systems with conflicting results. Perlmutter *et al.*, 1988, studied the evolution of the T15 family and concluded that the four V_H genes in this family diverged by up to 5% from their allelic counterparts. The suggestion was made that this divergence had occurred at a higher rate through frequent gene duplication and conversion events. Loh *et al.*, 1984, studied the sequences of five NP^b-related V_H gene segments from BALB/c and C57/B10 mice. They concluded that these genes had diverged so quickly that they were not able to define allelic relationships among them. Recently Kaartinen *et al.* (1989) have sequenced the V_H (Ox1) gene from several strains of mice and found a remarkably low level of variation. Four alleles of V_HOX1 have 99–99.7% sequence identity to each other. Our findings are most similar to these latter results.

Until recently only sporadic reports of V_H germline sequences have been reported to allow sequence comparison in order to assess the level of polymorphisms in human V_H genes. Yet it is striking that according to these reports at least some human V_H genes are maintained with 100% identity in unrelated individuals. This is, indeed, the case for some V_HIII genes including 20P1 which has been found expressed unmutated in the early human repertoire (Schroeder *et al.*, 1987) as well as in a human SLE specific anti-Sm autoantibody (Sanz *et al.*, 1989). There are also a few examples of expressed V_H genes and germline genes which are identical (Chen *et al.*, 1988; Pascual *et al.*, 1989). Recently, Van Dijk *et al.*, 1989, have used V_H specific oligonucleotides to detect minor sequence differences in human DNA. They were able to show, consistent with the results above, that the V_HIII family may display a high level of polymorphism in some family members whereas others like 20P1 are remarkably conserved. They also showed a lack of polymorphism in the V_HVI gene. Their conclusions are fully confirmed by our sequence analyses.

Here we report the sequence of 12 germline V_HIV genes from two unrelated Caucasian blood donors. We were able to assign allelic counterparts from donor V_H4.1 to most previously published sequences. Furthermore, in individuals that were heterozygous for some V_HIV genes we were able to identify both alleles by both sequencing and allele specific oligonucleotide probes. The conservation of this family is most remarkable. Out of eight cases where allelic relationships were studied, three were 100% identical, four range from 99–99.7% identity and one was 98%. Altogether only 14 differences were found in 2312 nucleotides sequenced which constitutes only 0.6% divergence. The ratio of transversions to transitions was 0.55 whereas on a random basis it should have been ~2.0, suggesting non-random divergence mechanisms. 78% of the mutations were located in the framework regions and 22% in the hypervariable region which is a random distribution. The frequency of replacement substitutions was 78% which is expected in a random distribution. 22% were silent mutations. Interestingly, all three mutations found in the hypervariable region were replacement substitutions yet the low number of mutations found prevents us from drawing statistically significant conclusions in this regard.

Using the silent substitution rate to measure evolutionary time, the divergence time for the V_HIV genes of the V_H4.1 donor would range from 0.5 to 1 million years (Kimura,

1981). All three parameters combined suggest that this family is diverging more slowly than expected. Furthermore, divergence mechanisms seem to occur randomly even though the suggestion can be made that evolutionary selection might be operating to fix transitions preferentialy over transversions and possibly to increase diversity in the hypervariable regions.

The V_HIV genes sequenced from donor V_H4.2 are significantly different. Only one (V_H4.23) was probably allelic to V.11 (99.3% identical) and to V_H4.17 (100% identical). The other two were 96.6% (V_H4.22) and 95.5% (V_H4.21) identical to V2-1 and V58 respectively. Therefore, at this point, we cannot conclude that we have found the V_H4.2 donor allelic counterparts of the reported V_HIV germline genes or of the V_H4.1 donor V_HIV genes. Rather we speculate that these genes might constitute a subfamily of V_HIV more closely related to V58 and V2-1 than to the rest of the family. Indeed, V58 and V2-1 are themselves the most distant members of the family when all the members are compared to one another. Furthermore, Lee *et al.*, 1987, have suggested that several V58-like genes might comigrate on an 18 kb *Hin*dIII genomic DNA fragment.

The analysis of V_HV and V_HVI genes is more straightforward. We have sequenced the allelic counterparts of the V_HV gene, V_H251, from eight unrelated individuals as well as the allelic connterparts of V_H32 from three unrelated individuals. The most distant allelic gene was 99.3% identical with only two substitutions, and five genes were 100% identical to V_H251 (corrected) and to one another. In much the same way, we sequenced the V_HVI genes from seven unrelated individuals and found that they were 100% identical to the three V_HVI genes reported in the literature and to one another. These data confirm and extend the findings of Van Dijk *et al.*, 1989, and suggest that this extraordinary conservation of V_HVI sequences may be of special importance as this gene is the most D_H proximal V_H gene segment and is expressed in the restricted fetal repertoire (Schroeder *et al.*, 1988).

Our data are consistent with several reports of identical nucleotide sequences found in expressed antibodies in unrelated individuals. This is in contradistinction with the level of polymorphism that has been shown or suggested for at least some members of human V_H families as well as for most (but not all, i.e. see Near *et al.*, 1984; Kaartinen *et al.*, 1989) reports on murine V_H genes. Interestingly, these conserved genes have been found expressed unmutated in both human and mouse autoantibodies (Sanz and Capra, 1988). Thus it is tempting to speculate that the ability to recognize self might constitute, at least for some genes, an evolutionary advantage which would work to fix these sequences in the human genome. Alternatively, it might be argued that these families might have evolved only recently and, therefore, have not had time for further diversification. Specifically, V_HV does not have a clear murine counterpart, the maximum homology being with some J558 members (67%). Yet the best homology found for V_HV among human V_H genes is with V_HI gens (the human counterpart of J558 genes) and ranges from 65 to 72%. This suggests that V_HV arose after speciation, underwent major changes in a brief period of time and has been remarkably conserved, probably due to some evolutionary advantage.

Materials and methods

DNA extraction

High mol. wt genomic DNA was extracted from granulocytes obtained from healthy blood donors as well as from healthy members of our laboratory. The normals and extended families analyzed were from various ethnic and racial populations while for cloning and sequencing all were Caucasian except V_HCHIN and V_HBLK which were derived from an Oriental and an African Black respectively. The extraction protocol was performed with an Applied Biosystems automatic DNA extractor (Foster City, CA).

Oligonucleotides

Oligonucleotide primers for polymerase chain reactions (PCR) were synthesized on an Applied Biosystems oligonucleotide synthesizer (Foster City, CA). The sequences of the oligonucleotides used for cloning V_HIV genes were V_HIV5′, 5′ CCT CTA GAC CCA GGG CTC ACT GTG GG 3′ [primes the V_HIV variable segment from the 5′ end (leader intron)] and V_HIV3′, 5′ CCC GGG CTC ACA CTC ACC TCC CCT 3′ [primes the V_HIV variable segment from the 3′ end (recombination spacer)]. These oligonucleotides prime all known V_HIV gene segments. The sequences of the oligonucleotides used for cloning the V_HV genes were 5′ GGG CCC TGA TTC AAA TTT TGT GTC TCC 3′ [primes the V_HV variable segment from the 5′ end (leader intron)] and 5′ GTC GAC GGG CTC GGG GCT GGT TTC TCT 3′ [primes the V_HV variable segment from the 3′ end (recombination spacer)]. These oligonucleotides prime V_H251 and V_H32 but do not prime the pseudogene V_H15. The sequences of the oligonucleotides used for cloning the V_HVI genes were 5′ CCC GGG AGA TGC CGT ATT CAC AGC AGC ATT 3′ [primes the single V_HVI gene segment from the 5′ end (leader intron)] and 5′ CCC GGG CTC ACA CTG ACT TCC CCT 3′ [primes the single V_HVI gene segment from the 3′ end (recombination spacer)]. Oligonucleotides used to screen colony lifts were as follows: (i) V_HIV internal: 5′ GCA GGT GAG GGA CAG GGT 3′ (FW1 AA17 → 22), (ii) V_HV internal: 5′ GGT AAA GCT GTA TCC AGA 3′ (FW1 AA25 → 30), and (iii) V_HVI internal: 5′ GCA ACA GTG CTG CTT GGA ACT 3′ (FW2 AA33 → 39).

Cloning and sequencing of germline V_H genes

Genomic DNA (1 μg) was amplified separately with the primers described above using the PCR in a Perkin Elmer DNA thermal cycler as described (Saiki *et al.*, 1986, 1988). Normally 40 cycles of amplification were performed, each cycle consisted of a denaturation step (94°C for 1 min), annealing for 3 min at a temperature 5°C below the calculated dissociation temperature for the specific primers used in each reaction, and an extension step (72°C for 3 min). At the end of the 40 cycles a further extension step was performed for 7 min at 72°C in order to increase the percentage of full-length blunt-end molecules. The PCR product was visualized in a 1% low melting point agarose gel and a fragment of the predicted size (~350 bp) was purified by phenol extraction and ethanol precipitation. The purified product was phosphorylated in the presence of 10 mM ATP with polynucleotide kinase and blunt-end ligated into a dephosphorylated, *Eco*RV digested p-Bluescript (KS + or KS−) vector or the *Sma*I site of pTZ18U. The ligation mixture was used to transform either BSJ72 or DH5 alpha F′ *Escherichia coli* competent cells. The colonies were screened with an internal oligonucleotide or a V_HIV cDNA (58P2) previously described (Schroeder *et al.*, 1987) labeled with [α-^{32}P]dCTP by random hexamer priming or with the V_HV and V_HVI internal oligonucleotides previously described. From the positive clones, single stranded template was prepared by superinfection with a M13 helper virus (VCS-M13, Stratagene). DNA sequencing was performed by the dideoxy chain termination protocol (Sanger *et al.*, 1980) using [^{35}S]α-thio-ATP and a modified T7 DNA polymerase (Tabor and Richardson, 1987).

Considerable caution was exercised to be certain that the same V_H gene segment was not being recloned (and/or reamplified) as a contaminant. Different phagmid vectors were employed, different individuals cloned and sequenced several of the genes independently; and different reagents, instruments and rooms were employed for some of the amplifications. Indeed, one amplification was performed off site to insure that we were not dealing with PCR artefacts.

Dot blot analysis using allele specific oligonucleotides

DNA was blotted onto MSI Magnagraph nylon membrane (MSI, Westboro, MA), baked for 1 h at 80°C, and prehybridized for 30 min at 42°C (6 × SSC, 0.5% SDS, 5 × Denhardt's solution and 100 μg/ml denatured salmon sperm DNA). The dot blots were probed in the same solution for 1 h at 42°C with probes that were designed to distinguish between putative

alleles in the V_HIV, V or VI families. An internal oligonucleotide, constructed from a region of each gene segment that showed no sequence variation, was used as an amplification control. The filters were then washed in a tetramethylammonium chloride (TMAC) solution (500 g TMAC, 70 ml 1 M Tris–HCl pH 8, 5.6 ml 0.5 M EDTA pH 8, and 7 ml 20% SDS) for 25 min at room temperature, then at varying temperatures depending on the oligonucleotide for 25 min. This removed any probe that had one or more mismatches with the target sequence. All samples were read independently by two authors. In addition, most dot blots were read in a Betascope (Betagen, Waltham, Massachusetts) radioactivity detector and quantitated.

Acknowledgements

The authors are grateful to Drs Virginia Pascual, Kathy Meek and Richard Baer for helpful discussions. We are grateful to Leon Carayannopoulos for his careful review of the manuscript. Margaret Wright and Angela Houston provided excellent secretarial assistance.

References

Baer,R., Chen,K.-C., Smith,S.D. and Rabbitts,T.H. (1985) *Cell*, **43**, 705–713.

Berman,J.E., Mellis,S.J., Pollock,R., Smith,C.L., Suh,H., Heinke,B., Kowal,C., Surti,U., Chess,L., Cantor,C.R. and Alt,F.W. (1988) *EMBO J.*, **7**, 727–738.

Bothwell,A.L.M., Paskin,M., Reth,M., Imanishi-Kari,T., Rajewsky,K. and Baltimore,D. (1981) *Cell*, **24**, 625–637.

Brodeur,Ph.D. and Riblet,R. (1984) *Eur. J. Immunol.*, **14**, 922–930.

Buluwela,L. and Rabbits,T.H. (1988) *Eur. J. Immunol*, **18**, 1843–1845.

Chen,P., Liu,M.-F., Sinha,S. and Carson,D.A. (1988) *Arthritis Rheum.*, **31**, 1429–1431.

Dersimonian,H., Schwartz,R.S., Barrett,K.J. and Stollar,B.D. (1987) *J. Immunol.*, **139**, 2496–2501.

Humphries,C.G., Shen,A., Kuziel,W.A., Capra,J.D., Blattner,F.R., and Tucker,P.W. (1988) *Nature*, **331**, 446–449.

Kaartinen,M., Solin,M.-L. and Makela,O. (1989) *EMBO J.*, **8**, 1743–1748.

Kimura,M. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 454–458.

Kodaira,M., Kinashi,T., Umerura,I., Matsuda,F., Noma,T., Ono,Y. and Honjo,T. (1986) *J. Mol. Biol.*, **190**, 529–541.

Lee,K.H., Matsuda,F., Kinashi,T., Kodaira,M. and Honjo,T. (1987) *J. Mol. Biol.*, **195**, 761–768.

Livant,D.C., Blatt,C. and Hood,L. (1986) *Cell*, **47**, 461–470.

Loh,D.Y., Bothwell,A.L.M., White-Scharf,M.E., Imanishi-Kari,T. and Baltimore,D. (1984) *Cell*, **33**, 85–93.

Makar,R., Sanz,I., Thomas,J.W. and Capra,J.D. (1988) *Ann. Inst. Past, (Immunol.)*, **139**, 651–657.

Matthyssens,G. and Rabbitts,T.H. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 6561–6565.

Near,R.I., Juszczak,E.C., Huang,S.Y., Sicari,S.A., Margolies,M.N. and Gefter,M.L. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2167–2171.

Pascual,V., Randen,I., Thompson,K., Forre,O., Natvig,J. and Capra,J.D. (1989) *J. Exp. Med.*, submitted.

Perlmutter,R.M., Berson,B., Griffin,J.A. and Hood,L. (1988) *J. Exp. Med.*, **162**, 1998–2016.

Rathbun,G.A., Capra,J.D. and Tucker,P.W. (1987) *EMBO J.*, **6**, 2931–2937.

Rechavi,G., Bienz,B., Ram,D., Ben-Neriah,Y., Cohen,J.P., Zakut,R. and Givol,D. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4405–4409.

Saiki,R.K., Bugawan,T.C., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1986) *Nature*, **324**, 163–165.

Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988) *Science*, **239**, 487–491.

Sanger,F., Carlson,A.R., Barrell,B.G., Smith,A.H.J. and Roe,B. (1980) *J. Mol. Biol.*, **143**, 161–657.

Sanz,I. and Capra,J.D. (1988) *J. Immunol.*, **140**, 3283–3285.

Sanz,I., Hwang,L.-Y., Hasemann,C., Thomas,J., Wasserman,R., Tucker,P. and Capra,J.D. (1988) *Ann. NY Acad. Sci.*, **546**, 133–141.

Sanz,I., Dang,H., Takei,M., Talal,N., and Capra,J.D. (1989) *J. Immunol.*, **142**, 883–887.

Schroeder,H.W., Hillson,J.L. and Perlmutter,R.M. (1987) *Science*, **283**, 791–793.

Schroeder,H.W.,Jr, Walter,M.A., Hofker,M.H., Ebens,A., Van Dijk,K.W., Liao,L.C., Cox,D.W., Milner,E.C.B. and Perlmutter,R.M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 8196–8200.

Shen,A., Humphries,C., Tucker,P. and Blattner,F. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8563–8567.

Siekevitz,M., Huang,S.Y. and Gefter,M.L. (1983) *Eur. J. Immunol.*, **13**, 123–131.

Tabor,S. and Richardson,C.C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4767–4771.

Tonegawa,S. (1983) *Nature*, **302**, 575–581.

Van Dijk,K.W., Schroeder,H.W.Jr, Perlmutter,R.M. and Milner,E.C.B. (1989) *J. Immunol.*, **142**, 2547–2554.

Received on August 14, 1989; revised on September 6, 1989