The EMBO Journal vol.7 no.3 pp.727–738, 1988

Content and organization of the human Ig V_H locus: definition of three new V_H families and linkage to the Ig C_H locus

Jeffrey E.Berman[1], Scott J.Mellis[1,5], Roberta Pollock[1], Cassandra L.Smith[2,4], Heikyung Suh[1], Bruce Heinke[1], Czeslawa Kowal[1], Urvashi Surti[6], Leonard Chess[5], Charles R.Cantor[3] and Frederick W.Alt[1,2]

The Howard Hughes Medical Institute and the Departments of [1]Biochemistry and Molecular Biophysics, [2]Microbiology, [3]Genetics and Development, [4]Psychiatry and [5]Medicine, College of Physicians and Surgeons of Columbia University, 701 W. 168th Street, New York, NY 10032, USA and [6]Department of Pathology, Magee Womens Hospital, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Communicated by B.Pernis

We present a detailed analysis of the content and organization of the human immunoglobulin V_H locus. Human V_H genes representing five distinct families were isolated, including novel members belonging to two out of three of the known V_H gene families (V_H1 and V_H3) as well as members of three new families (V_H4, V_H5, and V_H6). We report the nucleotide sequence of 21 novel human V_H genes, many of which belong to the three new V_H gene families. In addition, we provide a preliminary analysis of the organization of these gene segments over the full extent of the locus. We find that the five multi-segment families (V_H1-5) have members interspersed over nearly the full 1500–2000 kb of the V_H locus, and estimate that the entire heavy chain locus covers 2500 kb or less. Finally, we provide the first report of the physical linkage of the variable and constant loci of a human Ig gene family by demonstrating that the most proximal known human V_H segments lie within 100 kb of the constant region locus.

Key words: human antibodies / PFG / repertoire / V_H gene families/V_H gene organization

Introduction

Immunoglobulin (Ig) molecules are composed of a set of identical heavy (H) and identical light (L) polypeptide chains; amino-terminal variable regions of H and L chains interact to form the antigen binding domain. The variable regions consist of a framework of relatively conserved amino acid sequence interrupted by three highly variable ('hypervariable') regions referred to as complementarity-determining regions (CDRs) because they interact to form the antigen-contact area. H chain variable regions are encoded by three separate germline gene segments, variable (V_H), diversity (D), and joining (J) segments. L chain variable regions are encoded by V_L and J_L segments. Germline V_H and V_L gene segments each encode two CDRs; the third arises from the region where the gene segments are joined. Multiple segments of each type exist in the germline DNA of mammals; the occurrence of distinct germline variable region gene elements which encode different primary amino acid sequences is a significant source of antibody diversity. Diversity also arises somatically as a result of the gene assembly process and from somatic hypermutation of assembled genes. (Reviewed by Tonegawa, 1983; Kabat *et al.*, 1987; Berman *et al.*, 1987.)

The murine H chain locus has a cluster of four J_H segments just upstream from the first constant region gene; ~12 D segments lie from 1 to 80 kb upstream of the J_H cluster (Kurosawa and Tonegawa, 1982; Wood and Tonegawa, 1983; Tonegawa, 1983). Murine V_H gene segments lie at an unknown distance upstream of the D cluster; the exact number has been estimated from <100 to 1000s (see Discussion). V_H segments characterized thus far have been divided into nine families based on nucleotide sequence homology (Brodeur and Riblet, 1984; Winter *et al.*, 1985). In several murine strains, definition of families based on homology has been strengthened by findings that families are organized into relatively discrete clusters, although some interspersion has been observed (Kemp *et al.*, 1981; Brodeur and Riblet, 1984; Reth *et al.*, 1986; Rathbun *et al.*, 1987; Blankenstein and Krawinkel, 1987). With a few exceptions, members of different murine V_H families have not been physically linked (Reth *et al.*, 1986; Blankenstein and Krawinkel, 1987).

The J_H and constant region portions of the human V_H locus have an organization similar to that of mouse (Ravetch *et al.*, 1981; Flanagan and Rabbitts, 1982; Takahashi *et al.*, 1982; Max *et al.*, 1982; Shimizu and Honjo, 1984) and several human D segments have been isolated (Siebenlist *et al.*, 1982; Ravetch *et al.*, 1981; Bakhshi *et al.*, 1987; Berman *et al.*, in preparation). Human V regions were originally classified into three subgroups (I, II, III) based on amino acid sequences (Kabat *et al.*, 1979). Some human V_H gene segments have been identified based on homology with murine or previously isolated human V_H genes (Matthyssens and Rabbitts, 1980; Rechavi *et al.*, 1982, 1983) or from rearrangements in B cell malignancies (Takahashi *et al.*, 1984; Kudo *et al.*, 1985; Mensink *et al.*, 1986; Cleary *et al.*, 1986); these have been grouped into three families designated V_H1, V_H2 and V_H3 (corresponding to the three protein subgroups). In contrast to murine organization patterns, recent evidence suggested that members of the three known human V_H families are highly interspersed; a single cosmid clone contained representatives of all three families (Kodaira *et al.*, 1986). Further elucidation of the content and organization of the human V region locus and comparison to that of the mouse should provide fundamental insights into genetic mechanisms involved in the generation of normal and abnormal antibody repertoires (reviewed by Alt *et al.*, 1987).

©IRL Press Limited, Oxford, England

Table I. Human and murine V_H gene nucleic acid sequence homologies

	V_H1						V_H2	V_H3									V_H4		V_H5			V_H6
	21-2	1-1	7-2	22-1	15-1	8-2	CE-1	1-9 III	2-9 III	9-1	12-2	13-2	8-1B	15-2B	22-2B	2-3	1-9 II	2-9 II	5-1	5-2	1-V	6-1
(I)[c] V_H1 21-2	100																					
1-1	93	100																				
7-2	87	86	100																			
22-1	91	91	89	100																		
15-1[a]	91	91	89	94	100																	
8-2	89	86	81	82	82	100																
(II) V_H2 CE-1	54	53	54	52	51	53	100															
(III) V_H3 1-9III	70	67	64	68	67	68	59	100														
2-9III	61	61	57	58	60	63	54	78	100													
9-1	66	64	61	66	65	63	60	83	73	100												
12-2	67	66	64	67	67	64	56	83	74	89	100											
13-2	67	66	62	67	67	65	56	88	76	81	82	100										
8-1B	71	68	64	66	65	67	60	91	78	84	85	90	100									
15-2B	68	66	64	65	65	65	57	85	77	82	85	87	89	100								
22-2B	71	68	65	68	67	68	61	87	79	82	86	87	91	92	100							
2-3	66	65	63	69	67	62	54	79	73	87	89	79	82	82	83	100						
(II) V_H4 1-9II	65	63	60	62	60	60	66	67	62	65	65	64	68	66	68	63	100					
2-9II	65	63	60	61	60	61	64	67	61	66	66	65	69	65	68	64	94	100				
(I) V_H5 5-1R1	72	70	68	72	72	67	56	65	59	65	65	66	65	63	64	66	65	65	100			
5-2R1	71	70	68	70	70	67	55	65	59	65	65	64	65	64	65	65	64	65	95	100		
1-V	68	66	66	68	69	65	55	63	60	62	62	63	61	61	62	62	59	59	90	87	100	
(II) V_H6 6-1R1	56	54	53	56	55	57	59	64	60	61	60	61	63	62	62	58	70	70	56	55	56	100
(I)[d] 36-60	60	57	57	52	54	55	61	61	56	59	60	59	60	61	61	58	76	76	56	56	56	71
Q52	60	59	55	53	51	59	63	66	58	60	61	61	61	59	62	60	63	64	56	57	53	59
(II) B4(J558)	76	76	70	77	78	71[b]	51	63	54	60	58	59	60	59	59	59	60	60	67	67	63	56
(III) S107	63	62	60	64	62	61	57	74	65	76	78	75	76	76	76	78	62	61	63	61	62	58
J606	60	61	57	60	62	58	58	74	67	76	76	71	74	74	73	74	60	60	62	60	58	59
81X	65	63	62	60	60	62	58	78	65	72	73	77	79	75	77	71	61	62	60	59	55	57
441-4	64	63	59	62	60	60	60	74	69	74	74	72	75	74	76	71	65	65	61	61	58	60
31(3609)	52	52	50	49	49	53	71	57	50	55	56	56	57	56	59	54	61	61	53	53	53	57

Nucleic acid sequences of V_H coding regions (from codon −4 up to but not including the heptamer/nonamer or D segments) were compared and aligned for maximal homology using the Beckman MicroGenie sequence analysis program. The numbers represent percent homology values. All human V_H sequences shown were derived in the present study with the exception of V_HCE-1 (Takahashi *et al.*, 1984). V_H sequences representative of each murine family were obtained from the literature and except for V_H31 (Winter *et al.*, 1985) were derived from the same genes used to screen and isolate human V_H genes in the present study (see Materials and methods).

[a] V_H15-1 and 22-1 homologies represent comparisons up to the point of divergence at CDR2 and do not include the divergent 3′ end of these genes.

[b] Boxed homology values highlight the murine V_H gene family most closely related to each human V_H gene family.

[c] The protein subgroup classification of each human V_H gene family is shown in Roman numerals (Kabat *et al.*, 1979, 1987).

[d] The protein subgroup classification of each murine V_H gene family is shown in Roman numerals (Kabat *et al.*, 1979, 1987; Dildrop, 1984).

Results

Isolation of human V_H genes by homology to murine V_H genes

Human genomic DNA was digested with either *Eco*RI or *Hin*dIII and assayed by Southern blotting for hybridization to ^{32}P-labelled probes specific for seven of the nine known murine V_H gene families (J558, Q52, 36-60, S107, J606, 7183, 441-4). When assayed under 'normal' stringency hybridization conditions (50% formamide and 5 × SSC at 42°C), these probes hybridize only to members of homologous murine V_H families (Yancopoulos *et al.*, 1984, 1987). Previous work indicated members of the J558 and S107 families are homologous to human V_H families 1 and 3 (Rechavi *et al.*, 1982, 1983) and that a murine V_H not used in these analyses, 3609, is most related to the human V_H2 family (Lee *et al.*, 1987; Table I). Probes from murine families J558, S107, J606, 7183, and 441-4 (Figure 1A, lanes 1−8) identified human sequences, while the Q52 and 36-60 probes employed failed to hybridize to distinct sequences (see legend to Figure 1). The J558 probe (murine subgroup II; Kabat *et al.*, 1979) identified a distinct set of human DNA fragments (Figure 1A, lanes 7 and 8). The 441-4 and 7183 probes hybridized to a nearly identical set of fragments that were different from those detected by the J558 probe (Figure 1A, lanes 3−6); S107 and J606 probes hybridized to an identical subset of the fragments identified by 441-4 and 7183 (Figure 1, lanes 1−6). Although the latter four probes identify unique murine families, they are derived from genes which encode members of the same murine protein subgroup (III; Dildrop, 1984; Kabat *et al.*, 1987); apparently some individual human subgroup III genes retain regions of homology to more divergent murine subgroup III genes.

Murine probes which hybridized to human DNA fragments were used to screen a human *Mbo*I partial genomic library. Over 200 V_H-positive clones were obtained and 69 analyzed. All contained at least one V_H1 or V_H3 gene, but

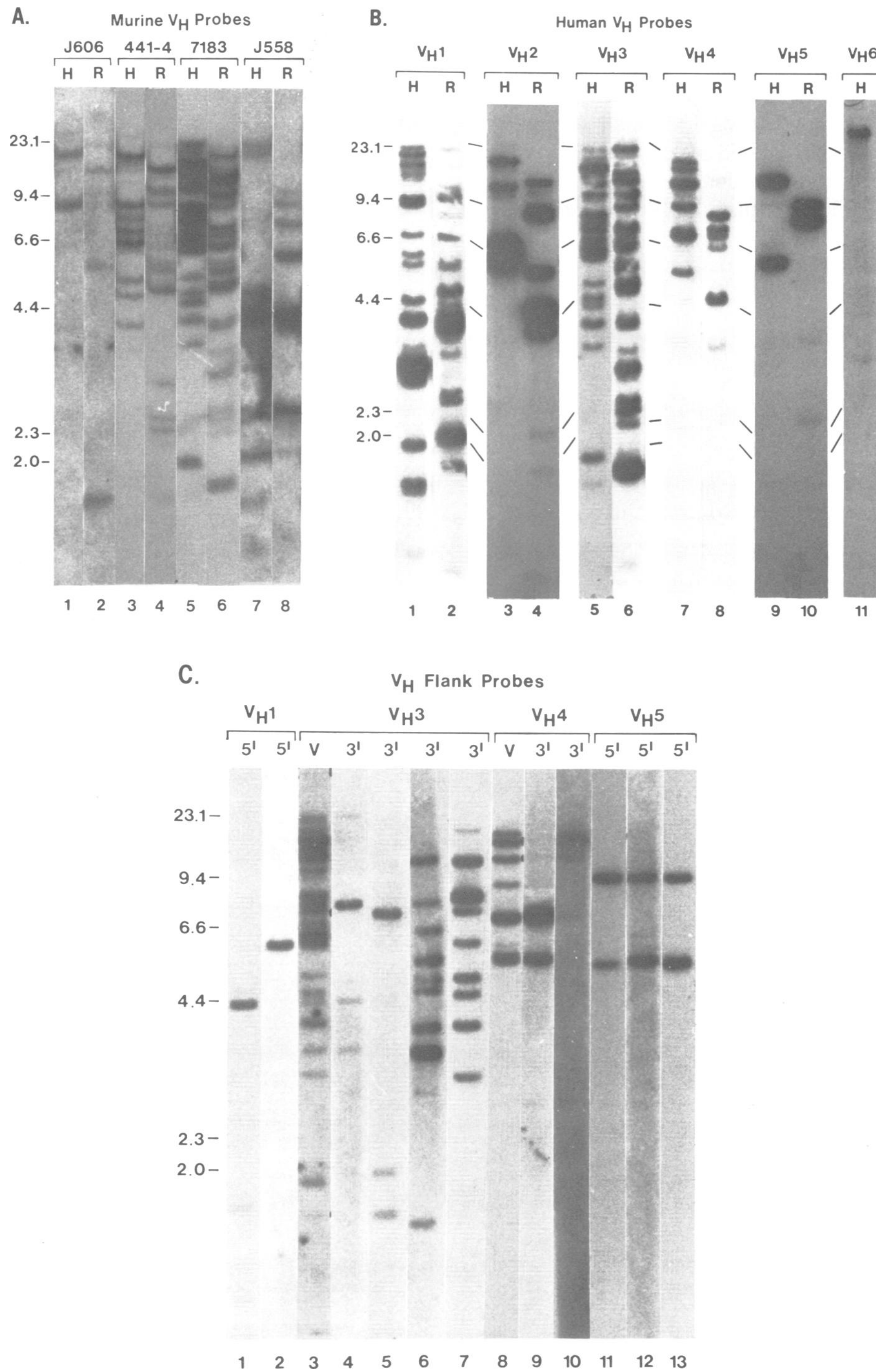


Fig. 1. Hybridization of human genomic DNA to murine and human V_H probes. Human or murine genomic DNA was cut with *Hin*dIII (H) or *Eco*RI (R), fractionated by agarose gel electrophoresis, transferred to nitrocellulose paper and assayed for hybridization to various ^{32}P-labelled V_H-specific probes. Sizes shown in kb were measured using *Hin*dIII-digested λ DNA. (**A**) Murine V_H probes. Genomic DNA was derived from the LA-N-5 human neuroblastoma cell line. The murine 36-60 probe employed was only 140 bp in length and spanned CDR2 (Yancopoulos *et al.*, 1987), which may explain why no hybridization to human sequences was observed (see text). (**B**) Human V_H probes. The various human probes are described in detail in Figure 3B. Human genomic DNA from either HeLa or LA-N-5 cell lines (both gave similar results) was assayed for hybridization with probes specific for the six known human V_H gene families: 15-1(V_H1), **lanes 1** and **2** [pattern obtained with 3-1(V_H1) was the same]; CE-1(V_H2), **lanes 3** and **4**; 9-1(V_H3), **lanes 5** and **6** (results with V_H3 probes, 12-2, 2-3, 9-1 and 22-2B were all very similar); 1-9II(V_H4)3′, **lanes 7** and **8** [1-9II(V_H4)5′ yielded the same result]; 5-1R1(V_H5), **lanes 9** and **10** [2-V(V_H5) yielded the same result but the relative intensity of the two bands is different]. The uppermost *Hin*dIII band corresponds to 5-1R1 and the lower *Hin*dIII fragment corresponds to 2-V (1-V), a pseudogene (see text). A third *Hin*dIII fragment is seen in some individuals and represents a polymorphism (P.Tucker, personal communication); 6-1R1(V_H6), **lane 11**, V_H6 is present on a 750 bp *Eco*RI fragment (data not shown; Figure 4D). Under slightly reduced stringency conditions this probe weakly cross-reacts with members of the V_H4 family. (**C**) Human V_H flank probes. *Hin*dIII-digested HeLa DNA was probed with the following probes; all lanes were derived from the same gel: 3-1(5′F), **lane 1**; 8-2(5′F), **lane 2**; 1-9III(V_H3), **lane 3**, coding region probe shown for comparison; 2-3(3′F), **lane 4**; 1-9III(3′F), **lane 5**; 12-2(3′F) **lane 6**; 22-2B(3′F), **lane 7**; 1-9II(V_H4)3′, **lane 8**, coding region probe shown for comparison; 1-9II(3′F), **lane 9**; 2-9II(3′F), **lane 10**; 5-1R1(5′F), **lane 11**; 2-V(5′F), **lane 12**; 2-V(5′F), **lane 13**. The family and the region of the V_H gene (5′ or 3′ flank or V_H coding region) from which the probes were derived is indicated (see also Figure 3B).

many bore more than one V_H gene—often from unrelated families (see below). Several clones bearing two V_H genes were mapped (Figure 3A), and the nucleotide sequence of V_H-positive restriction fragments from multiple clones determined. Two clones bearing V_H3 genes also contained another fragment which hybridized weakly to V_H3. These fragments were analyzed and found to contain V_H genes representing a new family, designated V_H4 (Figure 2D; see below).

Isolation of V_HDJ_H rearrangements from human lymphoid malignancies

Murine V_H probes hybridized to only two human V_H gene families under normal stringency conditions. Because some members of the V_H1 family (see below), the V_H2 family, and the newly identified V_H4 family were not detected by cross-hybridization, existence of additional V_H families seemed possible. To isolate V_H genes independent of murine homology, we cloned 11 V_HDJ_H rearrangements

A

B

C

D

from nine acute (ALL) and chronic (CLL) leukemia cell lines and tumors. Of the isolated rearrangements, four employed V_H3, two employed V_H1, and one employed V_H4 (Berman *et al.*, unpublished); the four which were negative for hybridization to known V_H probes contained members of two novel human V_H gene families, designated V_H5 and V_H6 (Figure 2; see below). The V_H portions of V_H5DJ_H and V_H6DJ_H rearrangements were used as probes to isolate germline phage clones bearing either V_H5 or V_H6 genes (Figures 2 and 4).

V_H gene families

The nucleotide sequences of 24 germline V_H genes and three V_HDJ_H rearrangements were determined; 21 were unique and none were identical to previously published V_H sequences (Figure 2A–D). The coding region sequences of the cloned genes indicated that they belonged to five distinct families based on definition of a family as a set of sequences sharing >80% homology (Brodeur and Riblet, 1984); members of different families had <70% homology (Table I). By comparison to the protein sequence classification of Kabat *et al.* (1979, 1987), the J558-homologous genes (e.g. V_H21-2) correspond to protein subgroup I (V_H1 gene family; Rechavi *et al.*, 1983), and the J606, S107, 7183, and 441-4-homologous genes to protein subgroup III (the V_H3 gene family; Matthyssens and Rabbitts, 1980; Rechavi *et al.*, 1982). The most divergent gene classified as a V_H3 sequence in this study (V_H2-9III) shared 73–79% homology with other V_H3 genes (Table I) but was a pseudogene due to multiple in-frame stop codons and abnormal recombination recognition sequences (Figure 2B). Thus, V_H2-9III appears to represent a divergent V_H3 pseudogene rather than a new family.

The two novel V_H genes (V_H1.9II and V_H2.9II) found on the phage clones with V_H3 genes when translated are classified as members of protein subgroup II. However, at the nucleic acid level, these genes are <70% homologous to any other known V_H families, even when compared to a previously reported V_H2 family gene, V_HCE-1 (Takahashi *et al.*, 1984), a member of subgroup II. Thus, protein subgroup II can be divided into two groups: V_H2 (e.g. V_HCE-1) and the other a novel V_H gene family, designated V_H4 (e.g. V_H1.9II and 2.9II; Table I; Malynn *et al.*, 1987). Lee *et al.* (1987) identified separate V_H4 family members.

The rearranged V_H genes cloned from two ALLs, as well as a corresponding germline V_H pseudogene (V_H1-V), shared 87–95% homology but were <70% related to V_H1–4 sequences; thus, these genes were classified as a new family (V_H5). Members of this family recently have been isolated by others from rearrangements in human CLLs (Shen *et al.*, 1987). Finally, the V_H segment employed in the rearrangement cloned from another ALL (6-1R1) was classified as a distinct family (V_H6) based on nucleotide sequence and Southern blotting analyses (see below); of the known V_H families, V_H6 is most related to V_H4 (70% homology; Table I). The 6-1R1 rearrangement was employed to isolate the corresponding germline V_H gene (6-1G1); 6-1G1 was identical in sequence to 6-1R1 except for the presence of several nucleotides deleted from 6-1R1 during the joining process (Figure 2D).

Of the 21 distinct V_H gene sequences presented here, eight were pseudogenes and the remainder appeared functional; at least one functional V_H gene was isolated from each family (except V_H2); the functional V_H5 genes coming from rearrangements. Characteristics of the various V_H genes including those for classifying genes as non-functional are in the legend to Figure 2.

Hybridization analysis of human V_H families

To further characterize the six human V_H families, ^{32}P-labelled probes specific for V_H segments representative of the five families isolated as described above (Figure 3B) and from a previously isolated V_H2 gene (CE-1, generously provided by T.Honjo) were assayed for hybridization to *Hin*dIII- or *Eco*RI-digested human genomic DNA. Under 'normal' stringency conditions these probes give six distinct patterns (Figure 1B, lanes 1–11), validating the conclusion that the V_H4, 5 and 6 genes truly represent novel families. The V_H3 probe identifies a very similar set of *Hin*dIII and *Eco*RI fragments (Figure 1B, lanes 5 and 6) to that detected with murine 441-4 and 7183 probes (Figure 1A, lanes 3–6). Thus, human V_H3 genes isolated using murine probes were fairly heterogeneous in sequence and appear to be a representative sampling of the entire V_H3 family (Table I and Figure 2). In contrast, the mouse J558 probe hybridized to a small subset of human V_H1 fragments (Figure 1A, lanes 7 and 8)

Fig. 2. Nucleotide sequence of human V_H genes. The nucleotide sequences of the V_H genes identified in this study are grouped into families based on >80% homology: V_H1 (**panel A**), V_H3 (**panel B**), V_H5 (**panel C**), V_H4 and V_H6 (**panel D**). Functional and pseudogenes are indicated by (+) or (−) signs, respectively. (R) indicates genes cloned as J_H-associated rearrangements, all of which were apparently formed from functional germline genes. Heptamer and nonamer recombination sequences are underlined. Previous studies have demonstrated that the size of the leader intron is common among members of a V_H gene family but varies between members of different families (Rechavi *et al.*, 1982, 1983). Thus, as expected, the novel V_H1 and V_H3 genes have intron sizes of 85 bp and 97–103 bp, respectively. The intron sizes of the other newly identified V_H families are slightly smaller, ranging in size from 82–83 bp in V_H4, 5 and 6. An exception is V_H2-9III (V_H3 family) which is a divergent pseudogene. The 6-1R1(V_H6) sequence is unusual in that it has a 20 amino acid leader, while most V_H genes have 19 amino acid leader sequences. Nucleotides belonging to a putative D segment in the 6-1R1 rearrangement are shown in lower case. The percentage of psuedogenes was not equally distributed among the different families; in particular only two out of nine different V_H3 genes were clearly pseudogenes. V_H genes were classified as pseudogenes for a number of reasons, including: lack of (e.g. 1-V) or defective (2-9III and 15-2B) heptamer/nonamer sequences, in-frame stop codons (e.g. 2-9II, 2-3 and 2-9III), and single base pair deletions resulting in frame-shift mutations (e.g. 1-1). V_H15-2B has an unusual splice site (AAGT) at the 3′ end of the leader intron and a C to G substitution in the third position of the highly conserved heptamer, but is otherwise normal. At least two genes (15-1 and 22-1) were truncated with completely divergent 3′ regions (see Discussion). The octamer sequence 'ATGCAAAT' (underlined) which has been found ~100 bp upstream of the V_H leader sequence and which has been implicated in the regulation of V_H gene expression (Parslow *et al.*, 1984; Falkner and Zachau, 1984; Grosschedl and Baltimore, 1985; Foster *et al.*, 1985) was identified without any changes in all eight of the V_H genes for which upstream sequence was derived, including examples of both functional and pseudogenes from V_H families 1, 3, 4, and 5. Sequences of V_H genes which proved identical are only shown once above (e.g. 15-2B = 17-2B, 7-1 = 21-2 = 3-1, 15-1 = 17-1 = 20-1, 1-V = 2-V).

Fig. 3. Human V_H gene clones. (**A**) Maps of phage clones. Partial restriction endonuclease maps of the inserts of human phage clones bearing more than one V_H gene are shown. Some of the germline V_H gene sequences presented in Figure 2 were derived from clones bearing only a single V_H gene, and thus are not shown. Only V_H genes which have actually been sequenced (Figure 2) are identified by name. Others were located by hybridization analysis and by analogy to overlapping clones. Family classification and transcriptional orientation (arrows) are shown below the name of each V_H gene. Functional and pseudogenes are indicated by (+) and (−) signs, respectively. In the case of 6-8-606 and 7-1-558, other overlapping phage clones are also shown below (within brackets). Restriction sites are as follows: H, *Hin*dIII; R, *Eco*RI; B, *Bam*HI; N, *Nco*I; X, *Xba*I. (**B**) Human V_H probes. V_H gene coding probes are denoted by stippled bars while V_H gene flank (F) probes are shown as striped bars. Germline V_H genes (shaded boxes) are shown with heptamer and nonamer recombination signals (open and closed triangles, respectively) except in cases where such sequences are lacking (e.g. 2-V). Rearranged V_H genes are depicted with D (open box) and J_H (stippled box) segments. The divergent portion at the 3′ end of V_H15-1 is indicated by a dotted box. The V_H2 probe (CE-1) was kindly supplied by T.Honjo (Takahashi *et al.*, 1984). Restriction enzyme sites are shown as follows: R, *Eco*RI;St, *Stu*I; P, *Pst*I, Pv, *Pvu*II; B, *Bam*HI; N, *Nco*I; X, *Xba*I; Hi, *Hinc*II; Rs, *Rsa*I; H, *Hin*dIII; S, *Sca*I; Hh, *Hha*I. Sites in italics are derived from the vector polylinker and are not found in the genome. The maps are drawn to the scale indicated except where distances are noted (in kb units).

identified by a human V_H1 probe (Figure 1B, lanes 1 and 2); thus, V_H1 clones isolated via J558 homology represent a subset of the V_H1 family (Table I). However, we were still able to isolate more divergent V_H1-representatives as the second members of clones bearing two V_H1 genes (e.g. V_H7-2 in clone 7-1-558; Figure 3A).

To facilitate mapping studies (see below), probes were derived from regions flanking various human V_H genes. Probes derived from the immediate 3′ flank of four V_H3 genes, two V_H4 genes, and from the 5′ flank of a V_H1 gene identified distinct nonoverlapping subsets of these respective families under normal stringency hybridization condi-

tions (Figure 1C, lanes 1, 3–7; 8–10). In contrast, probes derived from more distal 5′ regions of two V_H5 genes hybridized to the same set of fragments identified by coding sequence probes (Figure 1C, lanes 11–13; Figure 1A); nucleotide sequence analyses confirmed the conservation of 5′ flanking sequences among V_H5 genes (Figure 2C).

Organization of the human V_H locus

To examine V_H gene organization germline phage clones were analyzed with particular attention to those bearing two V_H genes. The location and orientation of V_H genes in phage inserts were mapped and in all cases found to be in the same transcriptional orientation (Figure 3A). V_H segments were separated by ~3.5–11 kb in these clones (Figure 3A), although analyses of other clones indicated that some segments are separated by larger distances (not shown). By screening human genomic libraries with the murine V_H probes, the following clones were obtained: 12 containing a single V_H1 gene, five with two V_H1 genes, 33 with single V_H3 genes; five with two V_H3 genes; 11 with both V_H1 and V_H3 genes, one with two V_H1 genes and a V_H3 gene and two with a V_H3 gene and a V_H4 gene. These findings support and extend the conclusions of Kodaira *et al.* (1986) that interspersion of V_H1, 2, and 3 members occurs at significant frequency.

To examine overall size of the H chain locus and extent of V_H gene family intermingling across the locus, genomic DNA from various cells was digested with *Not*I, fractionated by pulsed field gel electrophoresis (PFG), and assayed for hybridization to probes specific for human V_H families 1–6. Multiple sources of human DNA were employed including heterozygous cell lines, e.g. HeLa (Figure 4A, lanes 1–16), fibroblast (Figure 4A, lanes 17 and 18) and LA-N-5 (not shown) as well as a homozygous cell line derived from a hydatidaform mole (Figure 4B). In the heterozygous lines some *Not*I fragments could represent alternate alleles while the mole is an aberrant growth which bears a diploid homozygous genome and, as a result, is not subject to the uncertainties of genetic polymorphisms (Surti *et al.*, 1983; Surti and Szulman, 1986). Analyses of DNAs from all sources led to the same general conclusions. Thus, as with DNA from heterozygous lines (e.g. Figure 4A, lanes 1–16), all six V_H family probes hybridized to a common set or subset of *Not*I bands in mole DNA (Figure 4B). V_H1–4 hybridize strongly to a 670 kb fragment and a doublet at 570 kb (Figure 4B, lanes 1–4). V_H5 hybridizes strongly to the 570 kb doublet and very weakly to the 670 kb fragment (Figure 4B, lane 5); it is not clear whether the latter hybridization is due to the presence of a V_H5 gene or to minor cross-reactivity with another family. V_H6, a single member family hybridizes only to the 670 kb fragment (not shown, but see Figure 4A, lane 6). Two fragments show more restricted patterns: a 470 kb band is detected only by V_H1 and V_H4 probes and a 250 kb band only by V_H1 and V_H3 (Figure 4B). Together, these data indicate members of known multiple member V_H families (V_H1–5) are intermingled over much of the human V_H locus (summarized in Figure 4E).

To exclude possible misinterpretations that could arise as an artifact, for example, of incomplete digestion or cross-hybridization, we used V_H flank probes to confirm the unique nature of each *Not*I fragment. Because flank probes identify subsets of more closely related V_H genes within a family, these analyses also elucidated organization of V_H subsets. Several probes detected only single *Not*I bands demonstrating the digests were complete (Figure 4A, lanes 7–9; Figure 4B, lanes 6 and 9). For example, a 5′ V_H1 flank probe—which identifies a single *Hin*dIII fragment (Figure 1C, lane 2)—hybridized only to the smallest *Not*I fragment in HeLa (Figure 4A, lane 7), excluding the possibility that the three larger V_H-positive *Not*I fragments represent partially-digested precursors of the 250 kb fragment. Other V_H flanking probes identified various combinations of the V_H-positive *Not*I fragments (Figure 4A, lanes 10–12; Figure 4B, lanes 8, 10 and 11). The unique identity of the 670 kb *Not*I fragment in DNA from multiple sources also was confirmed by hybridization to the V_H6 probe and other probes from the Ig locus (see below). Thus, these results demonstrate that none of the *Not*I bands in HeLa or mole DNA represent partially-digested fragments.

Linkage of the human heavy chain variable and constant loci

To date, variable and constant Ig genes have not been physically linked in mammals. One study used Lod score analysis of restriction fragment length polymorphisms to estimate the V_H2 to J_H distance at 4 cM, which would represent roughly 4000 kb (Johnson *et al.*, 1984). To link human V_H, D, J_H and C loci more accurately, and to define J_H proximal V_H gene segments, additional PFG mapping experiments were performed. A J_H probe identified a 670 kb *Not*I fragment in both mole and HeLa which was the same size as the largest-sized V_H positive *Not*I fragment in the two lines (Figure 4, lanes A13 and 14, B12 and 13). The co-migration was confirmed by re-hybridizing the blots with a V_H3 probe (Figure 4, lanes A13 and 14, B12 and 13). However, it remained possible that J_H and V_H were present on two distinct but co-migrating *Not*I fragments. To address this issue, we incompletely digested human DNA with *Not*I to intentionally generate partially digested fragments and assayed the partial digests for hybridization with a V_H6 probe and with a J_H probe. V_H6 and J_H also identified a co-migrating *Not*I-partial fragment of 780 kb, confirming that these two loci occur on the single 670 kb *Not*I fragment (Figure 4A, lanes 17 and 18). In agreement with this conclusion, a D region-specific probe also hybridized to the V_H–J_H-positive, 670 kb *Not*I fragment (Figure 4A, lane 16). To map the H chain locus further, *Not*I-digested DNA was assayed for hybridization to C_μ and C_α probes; these probes again hybridized to the same 670 kb *Not*I fragment (Figure 4, lanes A15 and B14–16 and data not shown); the size of this fragment was conserved in all cell lines examined (e.g. Figure 4, lanes A13 and B12).

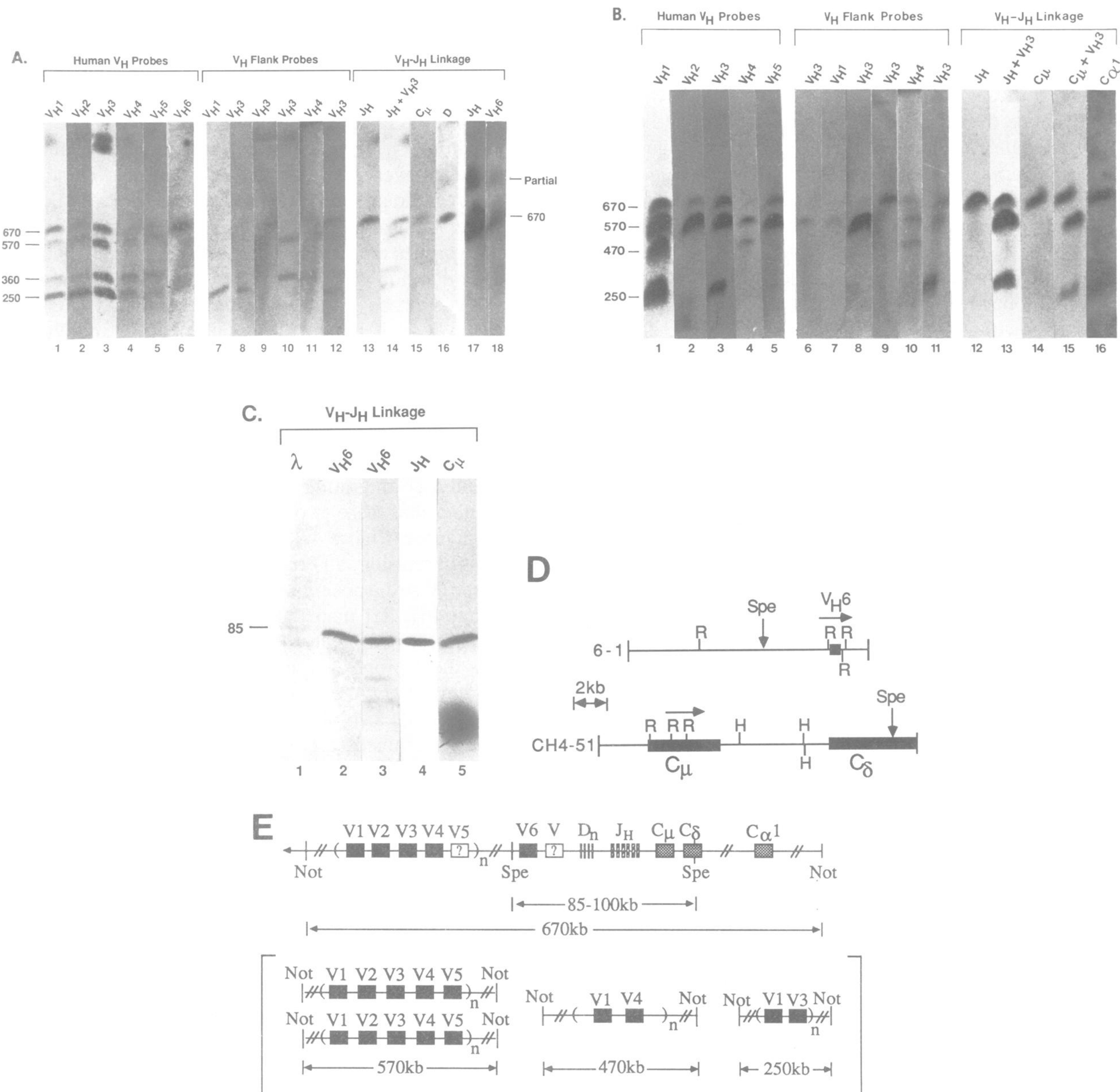
Additional analyses identical to those described above demonstrate that V_H6, D, J_H, and C_μ probes also hybridize to a common ~90 kb fragment generated by digestion with *Spe*I (Figure 4C). Analysis of incomplete digests, and double digests using *Spe*I–*Xho*I and *Spe*I–*Mlu*I (as outlined for *Not*I above) again confirmed the linkage of these sequences on the single 90 kb fragment; in addition, the *Xho*I–*Spe*I double digests generated two smaller fragments each containing D segments and either the V_H6 sequence or the C_μ region (data not shown). Thus, these data demonstrate that the distance between the proximal V_H segments and the J_H locus is significantly less than 100 kb (Figure 4D and E; see Discussion).

Discussion

Number of human V_H segments

Homologous counterparts of most human V_H gene families can be identified in the mouse (Rechavi *et al.*, 1982, 1983; Lee *et al.*, 1987). As outlined above and in Table I, V_H1 is homologous to J558; V_H2 to 3609; V_H3 to murine subgroup III (7183; J606; S107; 441-4); and V_H4 to 36-60. Murine counterparts of V_H5 and V_H6 are less obvious, but these families appear to be close relatives of human families 1 and 4, respectively. The relative importance of various factors which may have led to the similarity in V_H gene families between humans and mice are unclear. Similarities could be due to evolution from a common primordial V_H gene or to selection based on shared requirements for structure and diversity. As observed previously, the framework (FR) regions of the V_H genes within a family are considerably more conserved than CDR regions (Figure 2). Whether this phenomenon results from selective conservation of FR regions or an active mechanism which causes variability in CDRs is unclear (Kabat *et al.*, 1987).

Our current studies and those of others (see above) demonstrate that the human V_H gene locus is composed of at least six families. One method to estimate the total number of unique V_H genes in a genome is to add the total number of hybridizing *Hin*dIII or *Eco*RI restriction fragments which hybridize to probes specific for each family (Brodeur and Riblet, 1984). This estimate assumes an average distance between V_H genes of ~10 kb and therefore that an average sized *Eco*RI or *Hin*dIII fragment (theoretically 4096 bp) usually would bear only one V_H gene. By this approach, we estimate the following number of hybridizing *Hin*dIII (*Eco*RI) fragments (Figure 1B): V_H1, 20–25; V_H2, 5–10; V_H3, 25–30; V_H4, 6–10; V_H5, 2–3; and V_H6, 1, yielding a minimum estimate of ~60–80 unique V_H gene segments. However, there are limitations to this approach. Occurrence of coincidently migrating V_H-positive restriction fragments can result in an underestimate of V_H gene



numbers (Siekevitz *et al.*, 1983; Livant *et al.*, 1986). For example, V_H15-2B and V_H22-2B are distinct V_H3 genes and are present on co-migrating (8 kb) *Hin*dIII fragments, and V_H15-1 and V_H22-1, distinct members of the V_H1 family, are both found on *Hin*dIII fragments of 3.0 kb (Figures 2 and 3A). In addition, some V_H genes are more closely spaced than 10 kb (e.g. V_H7-1 and V_H7-2 are 3.5 kb apart; Figure 3A), again resulting in a potential underestimate of total number. Furthermore, the existence of additional, as yet uncharacterized, V_H gene families is possible. A separate method to estimate the number of genomic V_H segments is to determine the number of independent V_H-carrying clones in a single genome equivalent of a λ phage genomic library (Livant *et al.*, 1986). By this method we found 5-fold more V_H1- and V_H3-positive clones than expected based on Southern blot analyses (not shown), suggesting a total number of 300 or more. Of course, library screening procedures also have inherent difficulties such as relative cloning efficiency of different fragments. All results taken together, we consider it likely that the human V_H locus has between 100 and 200 members.

Our estimate for the number of human V_H genes is significantly less than that obtained for the total number of V_H genes in several murine strains (Livant *et al.*, 1986; Schiff *et al.*, 1985). In particular, the largest murine V_H gene family (J558) in some mouse strains has 50 or more hybridizing fragments and may have up to 1000 individual members (Livant *et al.*, 1986); the related human V_H1 family is not the most complex in humans and has about 20–25 hybridizing fragments. The sizes of other homologous families shared by humans and mice are more comparable. Thus, the total complexity (defined as the number of hybridizing *Hin*dIII or *Eco*RI restriction fragments) of the four families in mouse protein subgroup III is 28 (Brodeur and Riblet, 1984) while in humans the V_H3 family is composed of about 25 fragments; therefore, subgroup III appears similar in complexity in both species, although it may be somewhat less divergent in humans (see above). Similarly, the complexity of subgroup I in mice (Q52 and 36-60) is about 20 (Brodeur and Riblet, 1984) while the closest human equivalent, subgroup II (V_H families 2, 4 and 6) contains roughly the same number. Therefore, although the total number of V_H genes in the mouse may be greater than in humans, the total number of families is similar. Clearly, major differences in V_H gene numbers between mouse and human result from the size of the murine J558 family (subgroup II in mice). As there is no known difference in the extent of the diversity of the antibody repertoire between man and mouse, the size discrepency between the two germline V_H repertoires may result, at least in part, from functional redundancy within the murine J558 family. Of significance, the size of this family varies substantially among mouse strains (Blankenstein *et al.*, 1987).

Size of the human heavy chain locus

An approximate estimate of the total physical length of the human H chain locus (including V_H segments, D segments, J_H segments and the $C_\mu - C_\alpha$-1 region) can be made if one assumes that the V_H-positive *Not*I restriction fragments in the genomic DNA of the homozygous mole cell line can be arranged contiguously; with this qualification (see below) we estimate the size of the locus to be ~2500 kb (Figure 4B). We also can make a rough estimate of the size of the V_H-containing portion of the H chain locus by assuming that the overall size of the constant region portion of the locus is not drastically different from mouse (several hundred kb) and considering the possibility that the most 5′ V_H-positive *Not*I fragment only has V genes at its 3′ end; yielding a conservative estimate of 1500–2000 kb. Assuming an average spacing of 10 kb between V_H genes, this size is in rough agreement with that predicted by our estimate of the total V_H gene number.

We find that the V_H locus begins within less than 90 kb of the $J_H - C_\mu$ region. We do not know how close within this 90 kb the V_H and J_H segments lie. However, several pieces of information reflect on this. Firstly, we have identified *Spe*I sites in the immediate 5′ flanking region of the V_H6 gene and within the Cδ gene (Figure 4D). Thus, the maximum V_H to J_H distance must be less than 80 kb. Other data suggest that multiple D segments (Siebenlist *et al.*, 1982) lie within this region; it remains possible that at least one V_H gene may lie between V_H6 and J_H, although similar data

Fig. 4. Pulse field gel (PFG) analysis of the human V_H gene locus. HeLa cells (**panel A, lanes 1–16**), fibroblasts (A17 and 18), or hydatidaform mole cells (**panel B**) were immobilized in agarose blocks, lysed and their genomic DNA digested with *Not*I; subsequently the digested DNA was fractionated by PFG electrophoresis, transferred to nitrocellulose paper and assayed for hybridization to the indicated V_H, V_H-flank, D, or constant region probes. To demonstrate that fragments apparently detected by multiple V_H probes were the same, each probe was originally hybridized to a fresh PFG blot and then blots were stripped and re-hybridized to other family probes. Fragment sizes are indicated in kb units and were determined using λ concatemers and yeast chromosomes as standards (see Materials and methods). The probes used in each lane are as follows. Human V_H coding region probes: 15-1(V_H1), **lanes A1** and **B1**; CE-1(V_H2), **lanes A2** and **B2**; 9-1(V_H3), **lanes A3** and **B3**; 1.9II(V_H4), **lanes A4** and **B4**; 5-1R1(V_H5), **lanes A5** and **B5**; 6-1R1(V_H6), **lane A6**. Human V_H flanking region probes (the families from which they were derived are indicated above each lane): 8-2(5′F), **lane A7**; 12-2(3′F), **lanes A8** and **B6**; 3-1(5′F), **lane B7**; 2-3(3′F), **lanes A9** and **B9**; 1-9III(3′F), **lanes A10** and **B8**; 1-9II(3′F), **lanes A11** and **B10**; 22-2B(3′F), **lanes A12** and **B11**. D, J_H and C_H probes: J_H, **lanes A13** and **B12**; J_H probed and then reprobed without stripping with V_H3, **lanes A14** and **B13**; C_μ, **lanes A15** and **B14**; C_μ probed and then reprobed without stripping with V_H3, **lane B15** (the four V_H3-positive fragments correspond to those shown in the other lanes but were better resolved in this gel); Cα-1, **lane B16**; D, **lane A16**. **Panel A** (**lanes 17** and **18**) shows *Not*I partial digests of human fibroblast DNA (similar data from other lines not shown) fractionated by PFG [120 s pulse time, 10 V/cm, double inhomogeneous field, 55 × 55 cm chamber (LKB Produkter), 72 h run time]. To demonstrate co-migration blots were first probed with J_H (**lane A17**) and then stripped and reprobed with V_H6 (6-1R1, **lane A18**) or vice versa (data not shown). **Panel C** demonstrates linkage of V_H6, J_H and C_μ on a *Spe*I fragment of about 90 kb. *Spe*I digests of HeLa DNA were fractionated by PFG (pulse times were 50 s for lanes 1, 2 and 5 and 20 s for lanes 3, 4). Co-migration was shown by stripping and reprobing as described above [e.g. a J_H-probed blot (**lane 4**) was stripped and reprobed with V_H6 (**lane 3**) and vice versa (data not shown)]. Lambda concatemers (**lane 1**) were used to estimate the size of the V_H6-positive fragment (**lane 2**). The lower two faintly hybridizing bands in **lane 3** are due to cross-reaction of the V_H6 probe with members of the V_H4 family. For specific information on probes see Figure 3B (V_H probes) and Materials and methods (C_H and D probes). Co-hybridization of a uniqe 90 kb fragment to V_H6 and C_μ was confirmed by the use of partial digests as outlined in **panel A** and described in the text. **Panel D** shows the location of *Spe*I sites mapped in phage clones upstream of the germline V_H6 gene (6-1) and within the C_δ gene (CH4-51; Ravetch *et al.*, 1981); consistent with this, the C_δ sequence indicates the presence of a *Spe*I site in the C_δH2 exon (White *et al.*, 1985). Arrows above genes indicate transcriptional orientation. R, *Eco*RI; H, *Hin*dIII. **Panel E** represents a summary of the organization of the human V_H gene locus (in the mole cell line) as determined from the data presented above (**panels A–D**). $V_H - J_H$ linkage is shown in the upper half of the diagram. The lower portion of the diagram (in brackets) shows the other as yet unlinked V_H-positive *Not*I fragments including the 570 kb doublet (not drawn to scale).

map members of the V_H1, 3, and 5 families upstream of the 90 kb *SpeI* fragment (Figure 4E and data not shown).

Human V_H gene organization

V_H to DJ_H joining appears to occur predominantly by deletion in the mouse; indicating that V_H genes lie in the same transcriptional orientation as J_H segments (Cory *et al.*, 1980; Alt *et al.*, 1981; Tonegawa, 1983). All V_H genes that we analyzed were present in the same transcriptional orientation (Figure 3A), in agreement with previous V_H data from humans and mice (Kemp *et al.*, 1981; Givol *et al.*, 1981; Bothwell, 1984; Kodaira *et al.*, 1986). In fact, our present data linking the V_H6 gene and the J_H locus also demonstrates that these two segments are in direct orientation on the chromosome; so that joining would necessarily occur by deletion (Figure 4D). Results of deletion mapping indicate that many other human V_H gene segments also are oriented for deletional joining (R.Insel, J.E.Berman and F.W.Alt, unpublished results).

In the mouse, deletion mapping studies, studies of recombinant inbred strains and analysis of phage and cosmid clones containing multiple V_H genes suggested minimal interspersion of V_H families; that is, the various V_H families appear, in general, to be segregated as homologous families along the chromosome (Kemp *et al.*, 1981; Brodeur and Riblet, 1984; Brodeur *et al.*, 1984; Bothwell, 1984). Our data and that of others (Kodaira *et al.*, 1986) suggests members of different human V_H families are much more interspersed. Phage inserts containing two unrelated murine V_H genes rarely have been found; in contrast, we find that, for the largest families, individual phage inserts contain members of two different V_H gene families more frequently than two members of a single family (Figure 3A). In addition, PFG analyses suggest extensive intermingling of V_H gene families over the entire human locus. Although no definitive demonstration of intermingled functional V_H genes has been made, interspersion of functional V_H3 genes is suggested by the finding that V_H flank probes which define two functional V_H3 genes (Figure 2) hybridize to different *NotI* fragments (Figure 4A, lanes 8 and 10).

Widespread interspersion of divergent V_H genes may question a functional significance for clustering and ordering of particular V_H segments, although the human locus still could have evolved in the context of function. Recently, others have shown that human $V\kappa$ gene families (Pech and Zachau, 1984) and human T cell receptor V gene families are interspersed (L.Hood, personal communication). Extensive intermingling of divergent human V genes of a given type could be a random byproduct of the evolution of the locus or, theoretically, confer some selective advantage in the production of an antibody response as compared to a clustered organization. In mice V_H gene usage displays a highly position-related bias in fetal tissues relative to peripheral lymphoid organs in the adult, with J_H-proximal V_H genes utilized at a much higher frequency (Yancopoulos *et al.*, 1984; Malynn *et al.*, 1987; Alt *et al.*, 1987). Our current human map (Figure 4F) may support a similar phenomenon. Thus, the V_H6 family appears to be more highly represented in human fetal repertoires than those of the adult (J.E.Berman and F.W.Alt, in preparation), and a single member of the V_H3 family appears to be used at high frequency in fetal liver (Schroeder *et al.*, 1987).

Intermingling of V_H genes must be considered in theories of the generation of V gene diversity via amplification, contraction, and conversion mechanisms. Transposition events have been postulated in the evolution of human $V\kappa$ genes (Pech and Zachau, 1984). The data presented above suggest amplification, divergence, or transposition did not each occur once in simple sequence, but have instead occurred numerous times. For example, V_H3 gene flanking probes identified two distinct subsets of the V_H3 family each consisting of several members [one identified by 12-2(3′F) and one by 22-2B(3′F), Figure 1C, lanes 6 and 7]. Because the sequences flanking V_H genes tend to diverge, genes within subsets which share flanking homologies are presumably more closely related in evolution and each subset probably arose from amplification and divergence of a single ancestral V_H gene. When organization of these subsets was examined by PFG analysis, one [(12-2B(3′F)] was on a single *NotI* fragment while members of the other [22-2B(3′F)] were intermingled over the entire locus (Figure 4, lanes A8 and 12, B6 and 11). Thus, in the case of the 22-2B subset, more extensive transposition events appear to have occurred later in evolution after the emergence of an ancestral 22-2B gene, while in the 12-2 subset they have not.

It has been shown that some human $V\kappa$ genes are found on chromosomes other than 2, the location of the major body of the locus; the dispersed $V\kappa$ genes are pseudogenes and are usually more divergent than the pseudogenes found on chromosome 2, probably due to the absence of selective pressures (Lotscher *et al.*, 1986). Many germline V_H genes are pseudogenes. Although three of the *NotI* fragments clearly contain functional V_H genes (see above), it remains possible that one could bear divergent V_H genes at a dispersed chromosomal location. With respect to pseudogenes, two examples of highly related, but distinct, truncated V_H genes were isolated, V_H15-1 (V_H17-1) and V_H22-1; both appear normal and functional until the beginning of CDR2 at which point they diverge completely from V_H homology. The V_H22-1 sequence contains a normal upstream octanucleotide sequence; but both genes contain stop codons in all three reading frames, the first of which occurs near the point of divergence in the same reading frame as the coding sequence. Genomic Southern blots using a probe containing the entire V_H15-1 gene, including the divergent portion, show no obvious differences from that obtained with a functional V_H1 gene (3-1, data not shown); suggesting that these genes occur within the V_H locus. Similar examples of truncated but otherwise normal V_H genes have been described in chickens where they serve as substrates for gene conversion (Reynaud *et al.*, 1985). Although no strong evidence for gene conversion as a major somatic diversification mechanism has been obtained in mice, this issue remains to be addressed in humans.

Materials and methods

Cells

Human genomic DNA with an unrearranged Ig J_H locus was obtained from the following cell lines: HeLa, LA-N-5 (neuroblastoma; Kohl *et al.*, 1983), GM-3468 (normal fibroblast; NIGMS Cell Repository, Camden, NJ), HM811855 (hydatidiform mole). The mole cells were homozygous at the following loci by RFLP: H-*ras*, D11S12, insulin, γ G and γ A globulin, D14S1, α globin, and DXYS1. Chromosomal heteromorphisms examined were also homozygous (U.Surti, manuscript in preparation). J_H-associated rearrangements (V_H genes isolated from each sample are indicated in parentheses) were isolated from a variety of null ALL and CLL tumor samples

including the following: CAB (V_H5-1R1; null ALL PBL), BAK V_H6-1R1; pre-B ALL PBL), NALM-16 (V_H5-2R1; null ALL cell line). Murine genomic DNA was isolated from BALB/c liver.

Probes

The human V_H gene probes are described in Figure 3B. Murine V_H family probes are described by Yancopoulos *et al.* (1984, 1987). Other human probes used were the following: J_H (*Bam*HI−*Hin*dIII fragment containing the entire J_H region; Ravetch *et al.*, 1981), C_μ (0.9 and 1.1 kb *Eco*RI fragments from the genomic μ gene; Ravetch *et al.*, 1981), $C\alpha$-1 (2.1 kb *Sac* fragment from 9 kb 5′ of $C\alpha$-1; genomic clone kindly provided by P.Leder), and a 0.45 kb *NcoI*/*PstI* fragment located 0.8 kb from a D segment (J.E.Berman and F.W.Alt, unpublished results). Preparation of nick-translated probes and hybridizations were performed as previously described (Yancopoulos *et al.*, 1984).

DNA analysis

DNA preparation, restriction enzyme digestions, agarose gel electrophoresis, and blotting procedures were performed as previously described (Yancopoulos *et al.*, 1984).

Genomic cloning and DNA sequencing

The Charon 35 vector was kindly provided by Fred Blattner. Germline V_H genes were isolated by screening a Charon 35 recombinant phage library prepared from size-fractionated, *Mbo*I partially-digested DNA (Wilhelmine and Blattner, 1983) from the mole (in the case of V6-1G1) or LA-N-5 (all other germline sequences) cell lines with murine V_H gene probes (see above) or with human V_H probes 5-1R1(V_H5) and 6-1R1(V_H6) (see Figure 3B). Human J_H-associated rearrangements were isolated by screening phage libraries constructed from completely digested *Hin*dIII DNA cloned in Charon 21A (0−9 kb inserts; in the case of 6-1R1) or Charon 35 (9−21 kb inserts; in the case of 5-1R1 and 5-2R1), with a J_H probe (see above) as previously described (Mensink *et al.*, 1986). Most DJ_H rearrangements are on larger than 10 kb *Hin*dIII fragments (J.E.Berman and F.W.Alt, unpublished) and would not be isolated by this cloning procedure. Portions of the insert from phage clones were subcloned and the nucleotide sequence of relevant regions determined by the methods of Maxam and Gilbert (1980). DNA sequences were determined on both strands in most cases.

Pulsed field gel electrophoresis

DNA samples were prepared and digested in agarose plugs and gels were run as described by Smith and Cantor (1987). Except where otherwise noted in figure legends, electrophoresis was performed in LKB Pulsaphor units at 15°C, in 1 × TBE, 100 s pulse time, 10 V/cm, 40 h run time. DNA was transferred to nitrocellulose in 15 × SSC.

Acknowledgements

The authors would like to thank Gary Rathbun, Renate Dildrop, Richard Insel and Kathy Nickerson for critical evaluation of this manuscript. We would also like to thank Anna Shen and Fred Blattner, Phil Tucker, Roger Perlmutter and Tasuku Honjo for communication of their unpublished results. The authors would like to gratefully acknowledge the support of the Howard Hughes Medical Institute, grants from the NIH (AI-20047, CA-40427, CA-21112-09, GM14825, CA-43331, GM39782, AI-14969-10); DOE (DE-F602-87, ER-GD852); LKB Produkter AB, and an award from the Mallinckrodt Foundation (to F.A.). J.E.B. was a fellow of the Damon-Runyon/Walter Winchell Cancer Fund, S.J.M. a Pfizer fellow, R.P. a Special Fellow of the Leukemia Society, and F.W.A. an Irma T. Hirschl Career Scientist.

References

Alt,F.W., Rosenberg,N., Lewis,S., Thomas,E. and Baltimore,D. (1981) *Cell*, **27**, 381−390.

Alt,F.W., Blackwell,T.K. and Yancopoulos,G.D. (1987) *Science*, **238**, 1079−1087.

Bakhshi,A., Wright,J.J., Graninger,W., Seto,M., Owens,J., Cossman,J., Jensen,J.P., Goldman,P. and Korsmeyer,S.J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2396−2401.

Berman,J.E., Malynn,B.A., Blackwell,T.K. and Alt,F.W. (1987) *BioEssays*, **5**, 197−203.

Blankenstein,T., Bonhomme,F. and Krawinkel,U. (1987) *Immunogenetics*, **26**, 237−248.

Blankenstein,T. and Krawinkel,U. (1987) *Eur. J. Immunol.*, **17**, 1351−1357.

Bothwell,A.L.M. (1984) In Greene,M.I. and Nisonoff,A. (eds), *The Biology of Idiotypes*. Plenum Press, New York. pp. 19−34.

Brodeur,P. and Riblet,R. (1984) *Eur. J. Immunol.*, **14**, 922−930.

Brodeur,P., Thompson,M. and Riblett,R. (1984) *Regulation of the Immune System. UCLA Symposia on Molecular and Cellular Biology New Series*, **18**, 445−453.

Cleary,M.L., Meeker,T.C., Levy,S., Lee,E., Trela,M., Sklar,J. and Levy,R. (1986) *Cell*, **44**, 97−106.

Cory,S., Adams,J. and Kemp,D. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 4943−4947.

Dildrop,R. (1984) *Immunology Today*, **5**, 84−86.

Falkner,F.G. and Zachau,H.G. (1984) *Nature*, **310**, 71−74.

Flanagan,J.G. and Rabbitts,T.H. (1982) *Nature*, **300**, 709−713.

Foster,J., Stafford,J. and Queen,C. (1985) *Nature*, **315**, 423−425.

Givol,D., Zakut,R., Effron,K., Rechavi,G., Ram,D. and Cohen,J.B. (1981) *Nature*, **292**, 426−430.

Grosschedl,R. and Baltimore,D. (1985) *Cell*, **41**, 885−897.

Johnson,M.J., Natali,A.M., Cann,H.M., Honjo,T. and Cavalli-Sforza,L.L. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7840−7844.

Kabat,E.A., Wu,T.T. and Bilofsky,H. (1979) *Sequences of Immunoglobulin Chains*. NIH, Bethesda, MD, Publ. 80-2008.

Kabat,E.A., Wu,T.T., Reid-Miller,M., Perry,H.M. and Gottesman,K. (1987) *Sequences of Proteins of Immunological Interest*. U.S. Dept Health and Human Services, Washington, DC.

Kemp,D.J., Tyler,B., Bernard,O., Gough,N., Gerondakis,S., Adams,J.M. and Cory,S. (1981) *J. Mol. Appl. Genet.*, **1**, 245−261.

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

Kohl,N., Kanda,N., Schreck,R., Bruns,G., Latt,S., Gilbert,F. and Alt,F. (1983) *Cell*, **35**, 359−367.

Kudo,A., Ishihara,T., Nishimura,Y. and Watanabe,T. (1985) *Gene*, **33**, 181−189.

Kurosawa,Y. and Tonegawa,S. (1982) *J. Exp. Med.*, **155**, 201−208.

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

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

Lotscher,E., Grzeschik,K.-H., Bauer,H.G., Pohlenz,H.D., Straubinger,B. and Zachau,H.G. (1986) *Nature*, **320**, 456−458.

Malynn,B., Berman,J.E., Yancopoulos,G.D., Bona,C.A. and Alt,F.W. (1987) *Curr. Topics Microbiol. Immunol.*, **135**, 75−94.

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

Max,E.E., Battey,J., Ney,R., Kirsch,I.R. and Leder,P. (1982) *Cell*, **29**, 691−699.

Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499−560.

Mensink,E., Schuurman,R., Schot,J., Thompson,A. and Alt,F. (1986) *Eur. J. Immunol.*, **16**, 963−967.

Parslow,T.G., Blair,D.L., Murphy,W.J. and Granner,D.K. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2650−2654.

Pech,M. and Zachau,H.G. (1984) *Nucleic Acids Res.*, **12**, 9229−9236.

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

Ravetch,J.V., Siebenlist,U., Korsmeyer,S., Waldmann,T. and Leder,P. (1981) *Cell*, **27**, 583−591.

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

Rechavi,G., Ram,D., Glazer,R. and Givol,D. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 855−859.

Reth,M.G., Jackson,S. and Alt,F.W. (1986) *EMBO J.*, **5**, 2131−2138.

Reynaud,C.-A., Anquez,V., Grimal,H., Weill,J.-C. (1985) *Cell*, **40**, 283−292.

Schiff,C., Milili,M. and Fougereau,M. (1985) *EMBO J.*, **4**, 1225−1230.

Schroeder,H., Hillson,J.L. and Perlmutter,R. (1987) *Science*, **238**, 791−793.

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

Shimizu,A. and Honjo,T. (1984) *Cell*, **36**, 801−803.

Siebenlist,U., Ravetch,J.V., Korsmeyer,S., Waldmann,T. and Leder,P. (1982) *Nature*, **294**, 631−635.

Siekevitz,M., Huang,S.Y. and Gefter,M.L. (1983) *Eur. J. Immunol.*, **13**, 123−132.

Smith,C.L. and Cantor,C.R. (1987) *Methods Enzymol.*, **155**, 449−467.

Surti,U., Leppert,M. and White,R. (1983) *Am. J. Human Genet.*, **35**, 72A.

Surti,U. and Szulman,A.E. (1986) In Ichinoe,K. (ed.), *Sapporo International Seminar on Trophoblastic Disease*. Igaku-Sholn, Tokyo.

Takahashi,N., Ueda,S., Obata,M., Nikaido,T., Nakai,S. and Honjo,T. (1982) *Cell*, **29**, 671−679.

Takahashi,N., Noma,T. and Honjo,T. (1984) *Proc. Natl. Acad. Sci. USA*,

81, 5194–5198.
Tonegawa,S. (1983) *Nature*, **302**, 575–581.
White,M.B., Shen,A.L., Word,C.J., Tucker,P.W. and Blattner,F.R. (1985) *Science*, **228**, 733–737.
Wilhelmine,A.M.L. and Blattner,F.R. (1983) *Gene*, **26**, 171–179.
Winter,E., Radbruch,A. and Krawinkel,U. (1985) *EMBO J.*, **4**, 2861–2867.
Wood,C. and Tonegawa,S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3030–3034.
Yancopoulos,G.D., Desiderio,S.V., Paskind,M., Kearney,J.F., Baltimore, D. and Alt,F.W. (1984) *Nature*, **311**, 727–733.
Yancopoulos,G.D., Malynn,B.A. and Alt,F.W. (1987) submitted.

Received on November 10, 1987; revised on January 4, 1988