The EMBO Journal Vol.2 No.10 pp.1795 – 1800, 1983

Conservation and divergence of immunoglobulin V_H pseudogenes

Justus B.Cohen and David Givol*

Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot 76100, Israel

Communicated by D.Givol
Received on 4 July 1983

The 12 immunoglobulin V_H pseudogenes, that have been characterized to date, differ from most pseudogenes of other multigene families in two aspects: (i) they carry only one (11 cases) or at the most two (1 case) deleterious mutations and (ii) they show no evidence of increased divergence from intact V_H genes. We describe here the first immunoglobulin V_H pseudogene that does not have these characteristics. This pseudogene accumulated numerous deleterious mutations and diverged considerably from other genes of the V_H gene family to which it belongs. In possible contrast to the other V_H pseudogenes, this pseudogene seems to be selectively neutral. We discuss the implications of the characterization of this diverged V_H pseudogene in relation to our understanding of the genetic mechanisms that generate diversity among germline immunoglobulin V_H genes.

Key words: evolution/gene correction/Ig V_H pseudogenes

Introduction

Pseudogenes are believed to be evolutionary relics of once intact genes, which became selectively neutral with the occurrence of a first deleterious mutation. As a result, pseudogenes persist in the genome as freely drifting sequences equivalent to functionless flanking region or spacer DNA (for reviews, see Proudfoot, 1980; Little, 1982). These notions are based on two observations. (i) Different pseudogenes within a multigene family represent a wide range of divergence for the corresponding intact genes (Maniatis *et al.*, 1980). It is thus believed that their extent of divergence (and the number of inactivating mutations accumulated within them) is a measure of the time since their original inactivation. (ii) The types of mutations found in pseudogenes are similar to those characteristic of non-coding spacer DNA: in addition to point mutations (which are commonly found both between homologous coding and homologous flanking regions), pseudogenes often carry microinsertions and deletions (Efstratiadis *et al.*, 1980).

Large families of nearly identical genes (e.g., the snRNA genes, see Denison *et al.*, 1981) should provide a more favourable environment for pseudogenes than smaller multigene families, where the product of each gene has a distinct function (e.g., the globin gene family).

Immunoglobulin V_H genes seem to represent a multigene family with traits and requirements in common with both types of multigene families. V_H gene diversity is the basis of an efficient antibody response, so that a large and diverse V_H gene repertoire confers selective advantage. Inactivation of a single V_H gene may have a greater impact on selective fitness than inactivation of a single snRNA gene, because V_H genes are less numerous than snRNA genes, and their products serve non-identical functions (expressed as differences in specificity or affinity for antigen). On the other hand, V_H gene inactivation may be less harmful than inactivation of a certain globin gene, since V_H genes are more numerous, and the overlap in the function of their products tighter: possibly more than one V_H gene can contribute to the antibody response against a certain antigen. In addition, different products can result from a single V_H gene due to various somatic diversification processes (Tonegawa, 1983). Therefore, the extent of gene inactivation in a given multigene family may be a measure for the selective constraints imposed upon each gene of that family.

Previous studies from several laboratories, including ours, have indicated that the V_H gene repertoire is rich in pseudogenes (up to 40%, according to one estimation) (Bothwell *et al.*, 1981; Huang *et al.*, 1981; Givol *et al.*, 1981; Rechavi *et al.*, 1982, 1983; Loh *et al.*, 1983). One peculiar feature of all these pseudogenes (12 in total) is, that they carry only a single inactivating point mutation (11 cases), or, in one case, a point mutation and a 4-bp deletion (Table I). In addition, these genes still fall within the homology range found between intact V_H genes. These findings seem to indicate that these pseudogenes evolved only recently from normal genes. However, that may only reflect the fact that such pseudogenes are easier to detect precisely because of their high homology to intact genes. Perhaps more diverged pseudogenes may be found in the genome if one looks for them, so our previous estimation of 40% pseudogenes in the V_H gene repertoire may be too low.

Here we report the isolation and characterization of a truly diverged V_H pseudogene, which was identified because of its physical linkage to an active gene. Nucleotide sequence comparisons show that this pseudogene accumulated a relatively large number of deleterious mutations spread throughout the gene. Our results indicate that the V_H gene repertoire may contain many more pseudogenes than previously estimated. On the other hand, the high number of pseudogenes that did not diverge beyond one crippling mutation suggests that their sequences are being continuously corrected by processes such as gene conversion, within the multigene family.

Results

Loss of V_H genes during clone amplification

We have described the isolation of V_H gene-containing Charon 4A recombinant clones (Givol *et al.*, 1981). These clones were detected by screening a BALB/c embryo DNA library with a V_H cDNA probe, a *Pst*I fragment of plasmid pV(11)², which spans the region coding for amino acids 5 – 71 of the mouse plasmacytoma MPC11 V_H region (Zakut *et al.*, 1980). One of these clones, Ch124, gave a strong plaque hybridization, but upon isolation of its DNA only weak hybridization signals were detected with the V_H probe. Repurification of the plaque did not alter this finding, while an independent clone, which was isolated from a different BALB/c embryo DNA library, behaved in a similar way. Figure 1A (lanes 2 and 3) shows that the two clones yield

*To whom reprint requests should be sent.

Table I. Deleterious mutations in 12 V_H pseudogenes

Gene	Mutation	Position (codon no.)	Affected function	Homology		References
				with related V_H	between related intact V_H	
V_H145	TGC(cys) → AGC(ser)	22	protein folding	99.0% (V_H186.2)	95.1% (186.2/23)	Bothwell *et al.* (1981)
V_H104A	CAG(Gln) → TAG(ter)	39	terminator	90.5% (V_H105)	87.8% (105/108A)	Givol *et al.* (1981)
V_H3	TGG(trp) → TAG(ter)	36	terminator	92.1% (V_H33)	91.1% (104B[b]/33)	Loh *et al.* (1983)
V_H5	CAG/Gln) → TAG(ter)	1	terminator	87.2% (V_H33)	91.1% (104B[b]/33)	Loh *et al.* (1983)
V_H31	AGG(Arg) → TGA(ter)	50	terminator	85.6% (V_H33)	91.1% (104B[b]/33)	Loh *et al.* (1983)
V_H6	1-bp deletion	22	frameshift	96.7% (V_H186.2)	95.1% (186.2/23)	Bothwell *et al.* (1981)
V_H108B	ATG(Met) → ATA(Ile)	−19	initiation	91.5% (V_H108A)	87.8% (105/108A)	Givol *et al.* (1981)
V_H28	ATG(Met) → AAG(Lys)	−19	initiation	94.4% (V_H33)	91.1% (104B[b]/33)	Loh *et al.* (1983)
HA2	CAG(Gln) → TAG(ter)	6	terminator	88.9% (HG3)	–	Rechavi *et al.* (1983)
V_H111	TAT(Tyr) → TAA(ter)	94	terminator	86.9% (V_H105)	87.8% (105/108A)	Givol *et al.* (1981)
H16BR	CGA(Gly) → TGA(ter)	9	terminator	84.3% (H11)	–	Rechavi *et al.* (1982)
	CGC(Arg) → TGC(cys)	38	protein folding[a]			
	AGT(Ser) → TGT(cys)	53	protein folding[a]			
ψV3	TGC(Trp) → TAG(ter)	47	terminator	95.8% (V11)	90.2% (1/11)	Crews *et al.* (1981)
	4-bp insertion	90–91	frameshift			Huang *et al.* (1981)
	8-bp deletion	between 3′ signals	V-D-J joining			

[a]It is uncertain whether the indicated mutation is deleterious.
[b]The V_H104B sequence will be published elsewhere (Cohen and Givol, in preparation).

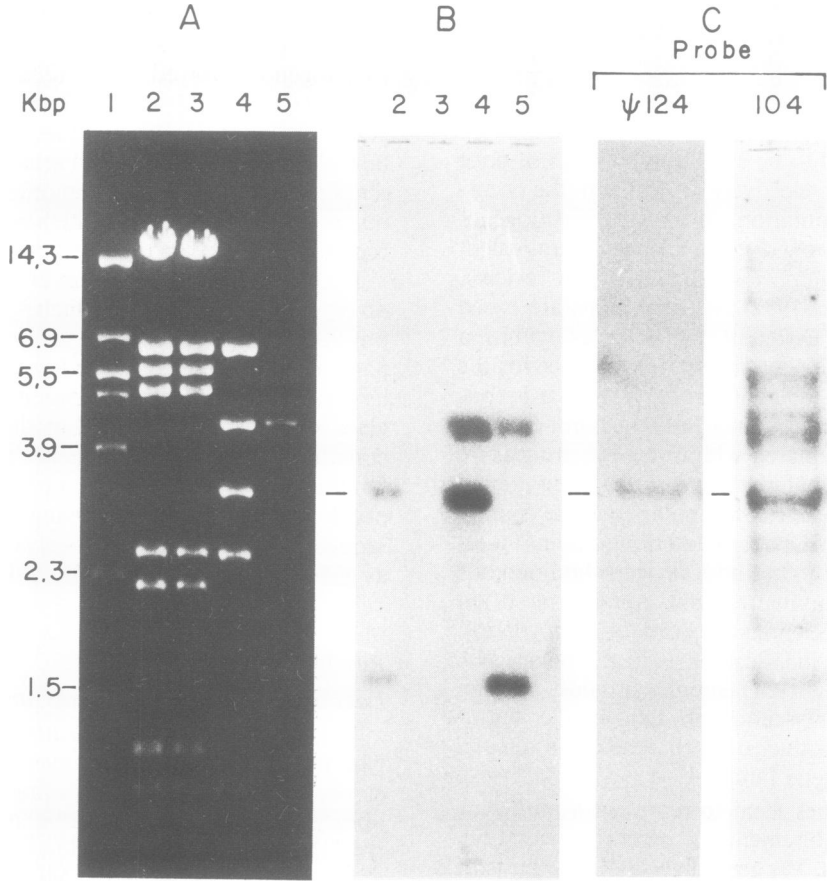

Fig. 1. Loss of V_H gene sequences during recombinant phage amplification and their recovery by subcloning (**A** and **B**), and evidence that pψ124 carries a diverged V_H gene (**C**). **(A)** *Hind*III-*Eco*RI digested Ch124 (**lane 2**), Ch2 (**lane 3**), pψ124 (**lane 4**) and p124 (**lane 5**) were electrophoresed on a 0.7% agarose gel with size markers (**lane 1**), and the fragments visualized by staining with ethidium bromide. DNA fragments were transferred to nitrocellulose paper according to Southern (1975), and the resulting blot hybridized to the V_H-containing *Bam*HI insert of plasmid pCh104 (Givol *et al.*, 1981). **(B)** Hybridization pattern of Ch124 (**lane 2**), Ch2 (**lane 3**), pψ124 (**lane 4**) and p124 (**lane 5**). **(C)** *Hind*III-*Eco*RI digested BALB/c liver DNA was electrophoresed in parallel on the same gel, blotted onto nitrocellulose filter paper, and hybridized to a probe prepared from the hybridizing fragment of pψ124, or to the *Bam*HI insert probe from pCH104. Washing conditions after hybridization were 3 x 30 min in 2 x SSC, 0.1% SDS, 0.1% $NaPP_i$ at 65°C (**lanes 2–5**), and 3 x 30 min in 1 x SSC, 0.1% SDS at 65°C followed by 30 min in 0.1 x SSC/0.1% SDS at 65°C (**C**).

identical restriction enzyme fragment patterns on agarose gels. When DNA of the second clone, Ch2, was analysed by Southern blot hybridization with various V_H probes, no signals were detected, while Ch124 DNA gave weak signals at positions where no ethidium bromide-stained bands were visible (Figure 1 B lanes 2 and 3). We concluded from these observations that the two clones carry one or two V_H genes, which are deleted at specific sites during amplification of the phage particles before DNA isolation. Since the Ch124 DNA still contained some of the presumably deleted fragments, as deduced from the hybridization signals on Southern blots, we decided to subclone total Ch124 digests in pBR322 in an attempt to obtain these genes in stable form.

Isolation of a diverged V_H gene

Subcloning of a *Hind*III-*Eco*RI digest of Ch124 DNA yielded both strongly and weakly hybridizing clones. Plasmids, isolated from the strongly hybridizing clones, contained an insert of 1.65 kb, and are called p124, whereas plasmids from the weakly hybridizing clones called pψ124, had an insert of 3.2 kb. Figure 1A (lanes 4 and 5) shows *Hind*III-*Eco*RI digests of both types of plasmid after electrophoresis in parallel with the bacteriophage digests in lanes 2 and 3. Comparison of lanes 4 and 5 with lane 2 in Figure 1B shows that each plasmid contains a hybridizing band that co-migrates with one of the hybridizing fragments of *Hind*III-*Eco*RI digested Ch124. In this figure the hybridization signal of pψ124 is as strong as that of p124, because the pψ124 lane contained much more DNA per fragment than the p124 lane, as can be seen from the picture of the ethidium bromide stained gel. The pψ124 subclone used in this case (Figure 1B lane 4) contains three additional *Hind*III-*Eco*RI fragments (apart from the pBR322 fragment), which all co-migrate with major bands in the Ch124 digest. These fragments are present in this subclone as a result of ligation (data not shown).

Plasmid pψ124 carries a diverged V_H gene

The weak hybridization of the putative V_H gene in plasmid pψ124 could be explained in two ways. First, this gene could belong to a V_H gene set which is only distantly related to the MPC11 V_H gene family (subgroup V_HII) that we have analysed so far. Alternatively, it could be a diverged member of this V_H gene family. To distinguish between these possibilities we hybridized parallel Southern blots of genomic BALB/c DNA to a probe that reveals the MPC11 V_H gene family (pCh104, Givol *et al.*, 1981), and to a probe prepared from plasmid pψ124. When we washed these blots under relatively mild conditions (1 x SSC, 0.1% SDS at 65°C for 3 x 30 min), we observed a similar pattern of bands with both probes, although some bands were weak or absent in the pψ124 probed blot. Figure 1C shows that, upon washing the filters after hybridization at higher stringency (i.e., an additional wash with 0.1 x SSC, 0.1% SDS at 65°C), only a single band remained in the pψ124 probed blot, whereas the pCh104 probe still detected a large number of fragments. The single band visualized by the pψ124 probe in Figure 1C co-migrates with the hybridizing *Eco*RI-*Hind*III fragment of pψ124 in Figure 1B, so that it is clear that the pψ124 probe detects only its own gene in BALB/c genomic DNA under stringent washing conditions. We concluded from this experiment that the gene of pψ124 is not a member of a distantly related V_H gene set, but that it is a diverged member of the V_H gene set which is detected by the pCH104 probe. We focused our attention on this gene to understand the reason for its divergence. The experiment of Figure 1B also shows that the gene, cloned in Ch124 and pψ124, faithfully represents this V_H gene in the BALB/c genome, and that it has not undergone gross rearrangement in the process of cloning, phage amplification, and subcloning.

Strategy for nucleotide sequence determination of the pψ124 V_H gene

Figure 2 shows how we determined the nucleotide sequence of the pψ124 gene. Sequences from the 5′ half of the gene were obtained by the procedure of Maxam and Gilbert (1980). The 3′ half of the gene, beginning at the single *Pst*I site at codons +4/+5, was subcloned into M13 vector mp8 (Messing and Vieira, 1982), and its sequence determined by the method of Sanger *et al.* (1980). The sequence of the 3′ half was also determined by the chemical degradation procedure after copying the single-stranded insert of M13 clones which contained this region (mp47), digestion with *Hind*III and *Eco*RI, and recloning into pBR322. The *Hind*III site from the polylinker region of M13 vector mp8 (Messing and Vieira, 1982), was now located next to the original *Pst*I site, so that sequence determination after end-labeling of the *Hind*III site included this *Pst*I site.

Plasmid pψ124 carries a V_H pseudogene

The nucleotide sequence of the hybridizing region in plasmid pψ124 was compared with the coding sequences of V_H genes that we previously isolated with the pV(11)[2] V_H fragment probe (Givol *et al.*, 1981; Cohen *et al.*, 1982). This comparison enabled us to locate the coding region of pψ124. Figure 3 shows the complete nucleotide sequence of this region in alignment with the previously described V_H gene 105 (Cohen *et al.*, 1982). Figure 3 reveals that the pψ124 gene is a pseudogene, which accumulated a relatively large number of deleterious mutations. We thus call this gene $\psi V_H 124$. In addition to crippling point mutations, such as those that alter the initiation codon at position −19 (ATG→AGG) and generate two termination codons in CDRII (TGA in positions 50 and 54), we find small insertions and deletions in the coding region of this gene. These mutations not only alter the translation reading frame of the gene, but also bring other

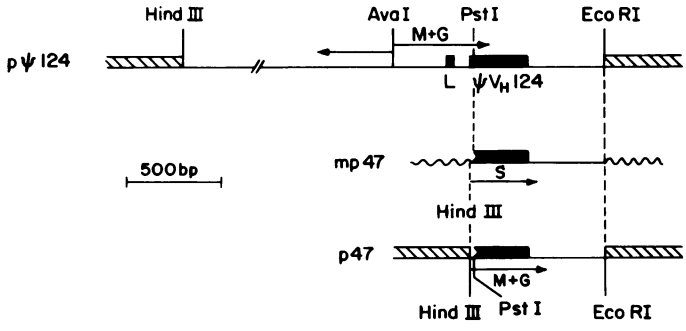

Fig. 2. Strategy used to determine the nucleotide sequence of the region in pψ124 that hybridizes to V_H probes. Black bars indicate regions that are homologous to the leader (L) and V_H gene segment of germline V_H genes. Hatched bars represent pBR322 vector sequences, and wavy lines M13 sequences. mp47 represents M13 mp8 recombinant phages which carry the *Pst*I-*Eco*RI fragment of pψ124 as indicated. p47 represents several pBR322 recombinant clones, which were obtained by cloning the insert of mp47 (including the *Hind*III site of the mp8 vector) back into pBR322. The procedure used for sequence determination in each case is symbolically indicated by 'M+G' for the chemical degradation procedure of Maxam and Gilbert (1980), and 'S' for the dideoxy chain termination procedure of Sanger *et al.* (1980).

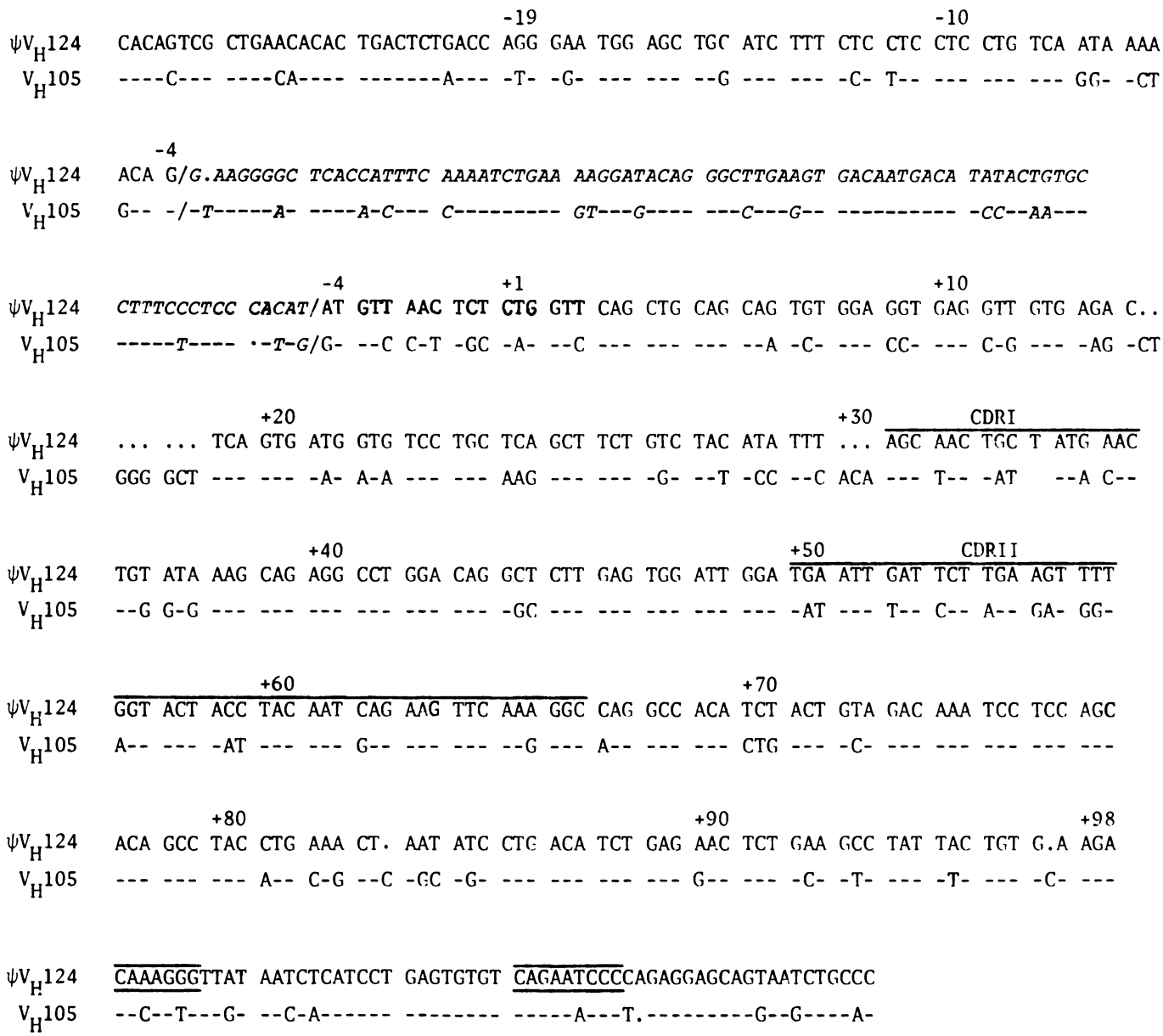

Fig. 3. The nucleotide sequence of ψV_H124 as compared with the previously published sequence of V_H105 (Cohen *et al.*, 1982). Sequences that align with the protein coding sequence of V_H105 are given in triplets, and the intron is in italics. Numbers above the codons refer to amino acid position according to Kabat *et al.* (1983). Complementarity determining regions (CDRI and II) are overlined. The signals for V_H-D joining are boxed. Dots within the sequence indicate deletions.

termination codons in phase with the altered reading frame (positions 29, 31, 35, 37, 72, 81, 83, 89, 92 and 97). In addition, several sequences which are believed to serve a DNA sequence-specific function in V_H genes, are mutated: the two dinucleotides at the borders of the intron (GA...AT) differ from the common GT...AG splicing signals (Breathnach and Chambon, 1981), and the first signal for V_H-D joining (CAAAGGG) differs at two positions from the consensus sequence (CACAGTG) (Tonegawa, 1983). As the latter signal is highly conserved between all V_H genes, and also between D genes, such changes may well interfere with V_H-D joining (Tonegawa, 1983). Finally, we note that ψV_H124 contains three additional codons for cysteine in the translation frame of a normal V_H gene (position 7, 33 and 36). Due to the grossly altered reading frame of the pseudogene, only the additional cysteine at position 7 is still relevant. However, we define a mutation as deleterious if that mutation by itself would have the potential of inactivating its gene, because in this way we list every mutation that could have originally inactivated the gene. Table II lists mutations in ψV_H124 that are deleterious by this definition.

ψV_H124 is selectively neutral

In Table I we have included information from the literature about the homology of the different V_H pseudogenes with intact genes of the V_H gene families to which they belong. These numbers are compared with homologies between two normal genes of the same family in each case, where such comparisons can be made. The data show that in all those cases the pseudogene did not diverge more than normal genes of the same family. We previously suggested that V_H pseudogenes are not free to diverge, as they may still serve some function (Givol *et al.*, 1981). For example, correction mechanisms like gene conversion could bring such pseudo-

genes back to the pool of intact V_H genes or make use of their sequence information to increase diversity among V_H genes (Baltimore, 1981). Hence, the maintenance of conserved V_H pseudogenes may be of selective advantage.

Pseudogene ψV_H124 shows a nucleotide sequence divergence of 23–26% from two of our previously analysed genes, V_H105 and V_H108B (Givol *et al.*, 1981), and 21.5% from the gene to which it is linked in the Ch124 clone, V_H124 (Cohen and Givol, in preparation). Genes V_H105, V_H108B and V_H124 represent different subsets of evolutionarily closely related genes within a single V_H gene family (Cohen *et al.*, 1982). Among the 12 genes of this family which have been sequenced so far (Bothwell *et al.*, 1981; Givol *et al.*, 1981; Cohen *et al.*, 1982; Loh *et al.*, 1983), the most diverged pairs differ by 16–20% in their V_H gene segment. ψV_H124 thus diverged from the majority of genes in this V_H gene family more than any pair of related genes. We conclude that this pseudogene must have escaped the forces that seem to restrict divergence of the other V_H pseudogenes.

We and others have previously shown that in the complementarity determining regions (CDRs) of germline V_H genes selection favours mutations that result in amino acid replacements (replacement substitutions), while it operates on the framework regions (FRs) to prevent such changes (Givol *et al.*, 1981; Loh *et al.*, 1983). This conclusion was based on the observation that the ratio of replacement (R) *versus* silent (S) substitutions is much higher in CDR than in FR regions. For example, Loh *et al.* (1983) recently reported a comparison of related V_H gene families from two different mouse strains, and showed that the overall value of R/S is 9-fold higher in the CDRs (7.0) than in the FRs (0.78). The values that we previously reported in a comparison of related V_H genes from a single mouse strain (Givol *et al.*, 1981) range from a factor of 6 (6.4/1.1) to 17 (20.3/1.2). Both Loh *et al.*, (1983) and ourselves (1983) attribute this phenomenon to strong selection on the protein level for conservation of the structure of the V_H region (determined by the FRs), and diversification of the antigen binding site (determined by the CDRs). In Table III we give the R/S values from a comparison of ψV_H124 with two intact genes of the same family, and compare these with the values obtained in our previous work. The contrast is striking: the R/S values in both comparisons of ψV_H124 differ by less than a factor of 2 between CDR and FR regions. These data clearly indicate that ψV_H124 has been free of selective pressure long enough to straighten out the enormous difference in R/S value between CRD and FR regions that is found between intact genes, and also between intact genes and other V_H pseudogenes.

Table II. Deleterious mutations in ψV_H124

Position (codon no.)	Consensus	ψV_H124	Affected function
−19	ATG	AGG	translation initiation
exon/intron borders	5′ AG/*GT*	AG/*GA*	RNA splicing
	3′ *AG*/GT	*AT*/AT	
7	TCT(ser,V_H105)	TGT(Cys)	protein folding[a]
14–16		8-bp deletion	translation reading frame
30		3-bp deletion	protein folding[a]
33	TGG(trp,V_H124)	TGC(Cys)	protein folding[a]
33–34		1-bp insertion	translation reading frame
36	TGG(trp,V_H124)	TGT(Cys)	protein folding[a]
50	GAG/glu.V_H124)	TGA	termination codon
54	AGA(arg,V_H105)	TGA	termination codon
83		1-bp deletion	translation reading frame
97		1-bp deletion	translation reading frame
3′ signals	CACAGTG	CAAAGGG	V_H-D-J_H recombination

[a]It is uncertain whether the indicated mutation is deleterious.

Discussion

With the availability of a considerable number of V_H gene sequences it is now evident that the V_H gene repertoire is rich in pseudogenes. One remarkable feature of V_H pseudogenes is that they do not differ greatly from what we consider to be normal, potentially active genes (Table I). Of the 12 V_H pseudogenes reported to date, 11 carry a single point mutation which presumably interferes with their expression. The twelfth gene carries an in-phase termination codon and a 4-bp insertion, which results in a shift of translation reading frame and the appearance of another in-phase termination codon (Huang *et al.*, 1981). Even this gene is still rather homologous (88–96%) to the different members of the T15 family of V_H genes, to which it belongs. These findings suggest that V_H pseudogenes are not free to drift in evolution, and we have previously proposed that some correction mechanism might operate to keep V_H genes and pseudogenes from drifting (Cohen *et al.*, 1982).

Because V_H gene families are relatively large, no attempt has been made to isolate and characterize all cross-hybridizing genes of a given V_H gene family. It is thus possible that truly diverged V_H pseudogenes were simply not reported because interest was generally focused on genes that are highly homologous to active genes. Only the slightly diverged pseudogenes would be selected in this way, while more diverged ones would not have been isolated. It is therefore possible that the number of V_H pseudogenes is larger than hitherto believed.

We decided to search for a diverged V_H pseudogene, so that we could decide which of the two possibilities mentioned above is responsible for the apparent lack of extensively drifted V_H pseudogenes. The identification of ψV_H124 shows

Table III. Ratios of replacement to silent substitutions (R/S) in pairs of V_H genes

Gene segment	R/S values					
	105/124[a]	108A/108B[b]	108A/111[b]	108A/104[b]	ψ124/124[a]	ψ124/105[a]
FR	0.54	1.0	1.1	1.2	2.1	2.3
CDR	4.00	11.6	6.4	20.3	3.5	4.0
CDR/FR	7.40	11.6	5.8	16.9	1.7	1.7

[a]The V_H105 sequence used for the comparison is from Cohen *et al.* (1982). The V_H124 sequence will be published elsewhere (Cohen and Givol, in preparation).
[b]Data are from Givol *et al.* (1981).

that the diverged V_H pseudogenes exist. This gene has accumulated a number of crippling mutations, and is apparently free of selective pressure. It was found in physical linkage to a normal gene, V_H124. The pseudogene is not more homologous to this neighboring gene than to other genes of this subgroup. Its presence in linkage with V_H124 indicates that diverged pseudogenes may be present between normal V_H genes. Some of these genes may have diverged so much that we will not detect them by means of cross-hybridization.

It is relevant to our understanding of the mechanisms that operate on V_H genes to create diversity without endangering their multiplicity to ask if diverged pseudogenes like ψV_H124 are indeed present in large numbers between V_H genes, or if ψV_H124 is an exception. The finding of many diverged V_H pseudogenes would indicate that V_H genes have a high turnover rate, so that new genes must be continuously created to keep a balance with the process of gene inactivation by drift. On the other hand, if ψV_H124 is an exception, and no evidence for large numbers of diverged V_H pseudogenes is found, then special mechanisms must be invoked to account for this lack of true pseudogenes and for the apparent accumulation of relatively large numbers of only slightly diverged pseudogenes like the 12 identified to date.

Southern blots of genomic mouse DNA indicate that the V_H gene family which we have analysed, consists of some 30 genes (e.g., Figure 1B), which hybridize rather strongly to our V_H probe (Ben-Neriah *et al.*, 1981). Hence, it is not possible to judge from such blots if diverged members of this family are present in the genome, and to what extent. We have therefore no basis to estimate which of the above mentioned alternatives is more likely. In work to be published elsewhere (Cohen and Givol, in preparation), we present evidence that gene V_H124, which is physically linked to ψV_H124, may have recombined in the past with a related V_H gene. The finding indicates that processes such as gene correction play a role in the evolution of V_H genes. Such processes could keep pseudogenes from drifting, and could possibly return them to the pool of intact genes. Further studies on isolated genes and particularly mapping of the entire locus by chromosome walking will eventually enable us to evaluate the role of such processes in V_H gene evolution.

Materials and methods

Subcloning into pBR322

5 μg pBR322 DNA was digested with *Hind*III and *Eco*RI, and the digest electrophoresed on a 0.3 mm thick 5% polyacrylamide gel (Sanger and Coulson, 1978). After staining of the gel with ethidium bromide, the larger fragment was electroeluted, and concentrated by ethanol precipitation. Approximately 10 ng of this fragment was ligated with T4 DNA ligase (New England Biolabs, MA) to 200 ng of *Hind*III-*Eco*RI digested Ch124 DNA. Transfection of HB101 bacteria was as described previously (Cohen *et al.*, 1982). Colonies of transformed bacteria were screened for the presence of V_H sequences by the method of Grunstein and Hogness (1975), using the V_H gene containing *Bam*HI fragment of plasmid pCH104 as a probe (Givol *et al.*, 1981).

Subcloning into M13 vector mp8

0.5 μg of the double-stranded (RF) form of M13 vector mp8 (Messing and Vieira, 1982) was digested with *Pst*I, *Bam*HI and *Eco*RI, and the digest ethanol precipitated. The *Bam*HI site is located between the *Pst*I site and the *Eco*RI site in the polylinker region of mp8 (Messing and Vieira, 1982), so that digestion with *Bam*HI splits the small fragment of *Pst*I-*Eco*RI digested mp8 into short oligonucleotides, which are presumably lost during ethanol precipitation. The digested vector was ligated to the isolated *Pst*I-*Eco*RI fragment of $p\psi124$, which carries the 3′ half of ψV_H124. Transfection of JM101 bacteria and screening by plaque hybridization have been described (Rechavi *et al.*, 1983).

Construction of plasmid p47

Plasmid p47 (see Figure 2) is a pBR322 recombinant, which carries the *Pst*I-*Eco*RI fragment of $p\psi124$ (containing the 3′ half of ψV_H124). This fragment was inserted between the *Eco*RI and the *Hind*III site of pBR322 by using a short stretch of the polylinker region in M13 recombinant clone mp47 as a *Hind*III linker, as follows. The single-stranded insert of mp47 was copied in the presence of the single-stranded M13 primer (Messing *et al.*, 1981) by the Klenow fragment of *Escherichia coli* DNA polymerase I (Boehringer, Mannheim, FRG), using [α-^{32}P]dATP (Amersham, UK) and an excess of unlabeled dCTP, dTTP and dGTP. After 15 min incubation at room temperature, an excess of unlabeled dATP was added, and the reaction was left to incubatefor another 15 min. The double-stranded insert was excised with *Hind*III and *Eco*RI, isolated by polyacrylamide gel electrophoresis and electroelution, and half of the material was used for cloning into the isolated large *Hind*III-*Eco*RI fragment of pBR322, as described above. The remaining half of the isolated radioactive mp47 insert was kept for use as a probe in the screening of transformed bacteria.

Nucleotide sequence determination

Nucleotide sequences were determined by the chemical degradation method (Maxam and Gilbert, 1980) as described previously (Givol *et al.*, 1981), or by the dideoxy chain termination procedure (Sanger *et al.*, 1980). Sequences of mp47 and p47 were determined on several independent isolates of these subclones.

Southern blot hybridization

Southern blots of restriction endonuclease digested cloned DNA or total BALB/c liver DNA were prepared as described by Southern (1975) and Rechavi *et al.* (1982). Hybridization and washing conditions were as described by Rechavi *et al.* (1982).

References

Baltimore,D. (1981) *Cell,* **24**, 592-594.

Ben-Neriah,Y., Cohen,J.B., Rechavi,G., Zakut,R. and Givol,D. (1981) *Eur. J. Immunol.,* **11**, 1017-1022.

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

Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.,* **50**, 349-383.

Cohen,J.B., Effron,K., Rechavi,G., Ben-Neriah,Y., Zakut,R. and Givol,D. (1982) *Nucleic Acids Res.,* **10**, 3353-3370.

Crews,S., Griffin,J., Huang,H., Calame,K. and Hood,L. (1981) *Cell,* **25**, 59-66.

Denison,R.A., Van Arsdell,S.W., Bernstein,L.B. and Weiner,A.M. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 810-814.

Efstratiadis,A., Pasokony,J.W., Maniatis,T., Lawn,R.M., O'Connell,C., Spritz,R.A., DeRiel,J.K., Forget,B.G., Weissman,S.M., Slightom,J.L., Blechl,A.E., Smithies,O., Baralle,F.E., Shoulders,C.C. and Proudfoot, N.J. (1980) *Cell,* **21**, 653-668.

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

Grunstein,M. and Hogness,D.S. (1975) *Proc. Natl. Acad. Sci. USA,* **72**, 3961-3965.

Huang,H., Crews,S. and Hood,L. (1981) *J. Mol. Appl. Genet.,* **1**, 93-101.

Kabat,E.A., Wu,T.T., Bilofsky,H., Reid-Miller,M. and Perry,H. (1983) *Sequences of Proteins of Immunological Interest,* published by NIH, Bethesda, MD.

Little,P.F.R. (1982) *Cell,* **28**, 683-684.

Loh,D.Y., Bothwell,A.L.M., White-Scharf,M.E., Imanishi-Kari,T. and Baltimore,D. (1983) *Cell,* **33**, 85-93.

Maniatis,T., Fritsch,E.F., Lauer,J. and Lawn,R.M. (1980) *Annu. Rev. Genet.,* **14**, 145-178.

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

Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.,* **9**, 309-321.

Messing,J. and Vieira,J. (1982) *Gene,* **19**, 269-276.

Proudfoot,N.J. (1980) *Nature,* **286**, 840-841.

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,L., Zakut,R. and Givol,D. (1983) *Proc. Natl. Acad. Sci. USA,* **80**, 855-859.

Sanger,F. and Coulson,A.R. (1978) *FEBS Lett.,* **87**, 107-110.

Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.,* **143**, 161-178.

Southern,E.M. (1975) *J. Mol. Biol.,* **98**, 503-517.

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

Zakut,R., Cohen,J.B. and Givol,D. (1980) *Nucleic Acids Res.,* **8**, 3591-3601.