Volume 12 Number 13 1984 Nucleic Acids Research

# Composite human $V_K$ genes and a model of their evolution

Hans-Rainer Jaenichen (1), Michael Pech, Werner Lindenmaier*, Norbert Wildgruber and Hans G.Zachau

Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München and *Gesellschaft für Biotechnologische Forschung mbH, Braunschweig-Stöckheim, FRG

Received 2 May 1984; Accepted 1 June 1984

ABSTRACT

A phage library and two cosmid libraries were screened for human $V_K$ genes. Two recombinant phage and four cosmid clones were analysed in detail by restriction mapping and sequencing. Each one contained a single $V_{KI}$ sequence. Two of these six sequences are potentially functional $V_K$ genes and four are pseudogenes. Two pseudogenes derived from different genomic DNAs are highly homologous and are therefore either allelic variants or the products of a recent duplication event. Comparisons of our sequences with all fully determined human $V_{KI}$ amino acid and DNA sequences reveal identical segments which at first sight appear like minigenes. But these segments do not coincide with the subregions and some of the segments include both, framework and complementarity determining regions (FR, CDR, ref. 2). The findings may be explained by an evolutionary model generating composite genes by gene conversion and selection.

## INTRODUCTION

The origin of antibody diversity used to be the subject of an intensive debate (3). The debate focussed on the nature and the origin of the variable regions of the antibody molecules. When a large number of amino acid sequences became available it was possible to delineate hypervariable regions within the variable part by a statistical analysis (4). These hypervariable regions are embedded in more conserved framework regions (FR). Hypervariable regions were the candidates for the molecular structures responsible for antigen binding and were therefore called *complementarity determining regions* (CDR). X-ray studies actually placed CDRs in close proximity to each other and confirmed that they could make contact with antigens (review ref.5).

In a compilation of all amino acid sequences then known and a comparison of their FR-sequences, E.A. Kabat et al. (6) found

© IRL Press Limited, Oxford, England.



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

sets of identical FRs. These identical subregions are shared by otherwise different polypeptide chains. The results were interpreted as an assembly of FR1-FR4 in which 'germline dictionaries of minigenes' (6) are used. The first sequences of germline $V_L$ genes (7,8) soon revealed that FR1 to CDR3 are encoded by a contiguous DNA sequence. Only the genetic information for FR4 is located on a separate minigene, the J-segment (review ref. 9). The question therefore arose how some DNA regions can be conserved during evolution even though they are linked to highly divergent segments in a multigene family (10). Recent reports on sequence homologies between D segments and CDRs of $V_H$ genes (11-13) revived a modified concept of minigenes as possible substrates for gene conversion-like events (11,14). The sequences of six $V_{KI}$ regions presented in this paper and their comparison with published $V_{KI}$ amino acid and DNA sequences support a model (14,15) in which gene conversion and selection upon sets of V genes may result in $V_K$ sequences that mimic assembled minigenes.

## MATERIALS AND METHODS

### Strains and materials

E. coli LE392 (supE supF $m_k^+ r_k^-$ met trp-R gal) was from P. Leder (16), E. coli 490A ($m_k^- r_k^-$ met thr leu $recA^-$) from G. Hobom (17), E. coli HB101 ($F^-$ pro leu thi lac Y $Str^r$ $m_k^- r_k^-$ Endo $I^-$ $recA^-$) from H.W. Boyer (18). E. coli JM103 ( $\Delta$(lacpro) supE thi straA end A sbc B15 hsd R4, F' traD36 proAB $lacI^q$, Z$\Delta$M15) originally constructed by J. Messing (19) was obtained from P.L. Biochemicals. DNAs of M13mp7, mp8, and mp9 (20) were also obtained from P.L. Biochemicals. The $\lambda$ phage library which was derived from HaeIII+AluI digested human fetal liver DNA inserted into Charon 4A (21) was kindly provided by T. Maniatis. [$\alpha$-$^{32}$P]dATP and [$\alpha$-$^{32}$P]dCTP used for nick-translation and sequencing were purchased from New England Nuclear, dideoxynucleoside triphosphates were from Boehringer Mannheim. The different M13 primers (15 bp, 17 bp, and 26 bp) used for the sequencing reactions were from P.L. Biochemicals or Bethesda Research Laboratories.

### Construction of cosmid libraries

High molecular weight placenta DNA (22) was used for the con-



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

struction of two cosmid libraries. The two cosmid vectors pHC 79-2 cos and pHC 79-2 cos/tk were previously described (23). For library I the DNA was partially digested with Msp I to yield fragments of about 40 kb and was then ligated to a fourfold molar excess of Cla I cleaved pHC 79-2 cos. Total DNA concentration during ligation was about 500 µg/ml. Library II was constructed using the same placenta DNA cleaved partially with Msp I, which was then separated on a 5 % to 20 % NaCl density gradient (24). Fractions containing DNA fragments from 25 to 50 kb lengths were pooled and ligated to a fourfold molar excess of Cla I cleaved and dephosphorylated pHC 79-2 cos/tk. Total DNA concentration during ligation was about 700 µg/ml.

In vitro packaging of both cosmid libraries was done as described (25). For the amplification of cosmid library II the packaged cosmids were divided into twelve fractions and used for transduction into $\lambda$-lysogens as described (23). The twelve pools of ampicillin resistant colonies consisting of $2.5 \times 10^4$ clones each were resuspended in L-broth to an $OD_{600} = 0.1$, grown at 30° C to $OD_{600} = 0.3$ and induced for in vivo packaging as described (23). The lysates contained about $1 \times 10^7$ transducing particles per ml. The packaged cosmids of library II were used for transduction into E. coli 490A.

Screening of cosmid libraries

Cosmid containing bacteria (E. coli HB 101 for library I, E. coli 490A for library II) were plated directly onto nitrocellulose filters placed on L-agar plates (150 mm Ø) at a density of $5\text{-}10 \times 10^3$ colonies/plate. A total of $2.5 \times 10^5$ colonies of library I and about $3 \times 10^5$ colonies of library II were screened according to ref. 26. The filters were hybridized to a nick translated 3.5 kb BamHI fragment from the M13 subclone 1-1 (Fig. 1); hybridization conditions were 6xSSC, 0.1 % SDS, 10 x Denhardt's solution (27) at 68° C for 12 hrs. with a final washing step in 3xSSC at 68° C. Colonies yielding positive signals with the human $V_K$ probe were identified on the master plate, isolated and purified by two additional rounds of screening.

Screening of the Ch4A phage library

Recombinant phages of the Ch4A library prepared by R.M. Lawn et al. (21) were amplified according to ref. 28 and plated on



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

E. coli LE392. A total of $2x10^6$ plaques at a density of $3.5x10^4$ plaques/plate (150 mm Ø) was screened by the plaque hybridization assay (29) using the pBR322 subclone L6/5 (30) as a probe which was derived from a mouse $V_K$ germline gene region. Hybridization conditions were the same as described for the screening of cosmid libraries.

Sequence analysis

DNA regions crosshybridizing to L6/5 or the M13 subclone 1-3 (Fig. 1) were sequenced according to the chain termination method (31) using two different sequencing strategies. Most sequences were determined on restriction fragments cloned in M13 phages. The fragments were selected according to their hybridization with specific subclones and on the basis of detailed restriction maps of the $V_K$ gene regions. The second strategy which was applied to sequencing most of V1 and V2 involved a shotgun procedure similar to the one of ref. 32. Instead of DNAase I digestion (32) sonication of DNA was used for generating a set of DNA fragments which were size fractionated, treated with E. coli DNA polymerase I (large fragment), and ligated into M13mp7 which had been linearised by HincII. Sequences were then completed by cloning and sequencing of defined restriction fragments. For determination of long sequences the sequence reaction products were electrophoresed on 6 % polyacrylamide / 7M urea gels (100x20x0.025 cm) which were thermostated (33).

Computer programs

The programs of R. Staden (34) were used for assembling and editing of the sequences. The programs DNMAHO, DNPERC, and DNCHOD are described in ref. 15.

## RESULTS AND DISCUSSION

### Screening for human $V_K$ genes with the help of a mouse $V_K$ probe

The phage library derived from human fetal liver DNA (21) was screened with a mouse germline $V_K$ gene probe (Materials and Methods). Plaques yielding intensive hybridization signals were purified and used for the preparation of recombinant phage DNA. Three of the clones (Ch1, Ch2, Ch4) were selected for further analysis. The restriction maps of Ch1 and Ch2 are presented in Fig. 1. Various subclones of Ch1 were constructed and used for



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

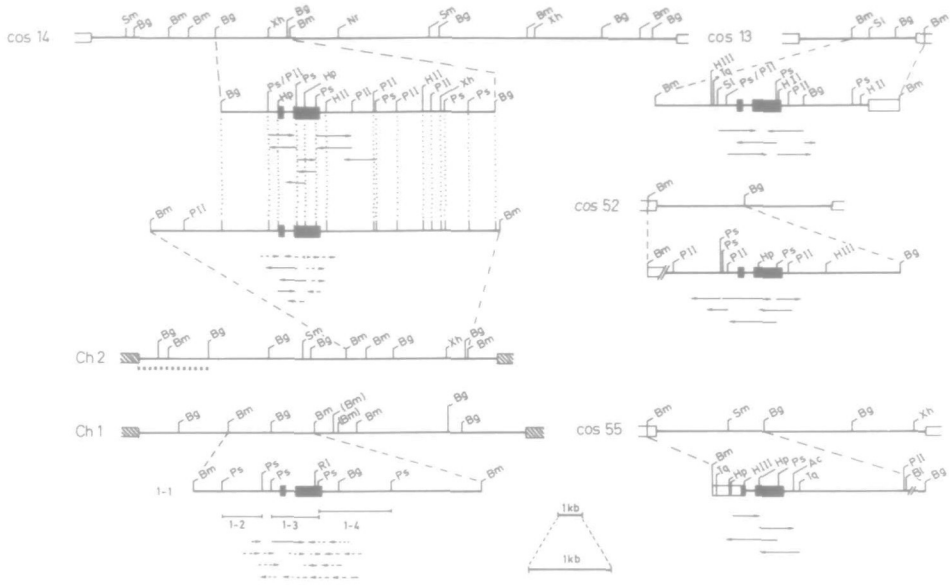



Figure 1. Restriction maps and sequencing strategies of six human $V_K$ gene regions. Restriction maps of four recombinant cosmid and two phage clones are shown. Fragments containing $V_K$ gene regions were further analysed by detailed restriction mapping (extended maps). The dotted vertical lines connect identically spaced restriction sites in the expanded maps of cos 14 and Ch2 in order to demonstrate the extent of homology between the two $V_K$ regions. Abbreviations for clones and subclones (underlined in the expanded map of Ch1) are described in ref. 2. Nucleases are abbreviated: Ac, AccI; Bm, BamHI; Bg, BglII; HII, HincII; HIII, HindIII; Hp, HpaII; Nr, NruI; Ps, PstI; PII, PvuII; RI, EcoRI; Sl, SalI; Sm, SmaI; Tq, TaqI; Xh, XhoI.

The extent and direction of sequencing are symbolized by contiguous arrows for the sequencing strategy employing cloned restriction fragments and by broken arrows for the shotgun approach (Methods). Closed boxes indicate the segments coding for the L and V regions; open boxes symbolize the cosmid and hatched boxes the phage vector DNA. Two additional BglII sites within Ch2 could not be mapped unequivocally and are omitted within the region underlined by dots.

screening and analysing recombinant cosmid clones. The M13 subclone 1-1 (Fig. 1) contains about 3.0 kb of flanking sequences in addition to the $V_K$ gene segment. M13 subclone 1-2 is a 5' flanking probe and subclone 1-4 comprises 54 bp of FR3 and CDR3 and about 830 bp of the 3' flanking region. The restriction map and sequence analysis of Ch4 is described in ref. 15.



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

Construction of cosmid libraries and analysis of four recombinant cosmid clones

Two cosmid libraries were prepared according to different experimental protocols (Materials and Methods). For the construction of both libraries (I and II) MspI partial digests of the same human placenta DNA were used. The insert of the M13 subclone 1-1 was used to screen both cosmid libraries. A total number of 27 recombinant cosmid clones were isolated and analysed by detailed restriction mapping. Only 15 of these clones yielded a positive hybridization signal with the smaller $V_K$ gene probe 1-3 (Fig. 1). This subclone contains the part of the V gene segment coding for amino acids 1 to 79, the leader, the intron, and 125 bp of the 5' flanking region. Four of these 15 cosmids (cos 14, cos 13, cos 52, cos 55) which hybridized with the 1-3 probe were further analysed and their restriction maps are shown in Fig. 1; one other recombinant cosmid is described in ref. 15; most of the remaining ones have not been analysed much beyond restriction mapping. An analysis of cosmid clones which did not contain a $V_K$ gene is presented in the accompanying paper (35).

According to the restriction maps of the isolated recombinant cosmids which hybridized with the 1-1 probe about half of the clones in both libraries have more than one copy of the cosmid vector and one or two inserts of placenta DNA. Several inserts were only about 10 kb long or even smaller (Fig. 1). This may be due to a lack of an efficient size fractionation of the MspI digests and/or to an incomplete phosphatase treatment in the preparation of the cosmid libraries.

One of the four cosmid clones presented in this paper, cos 14, contains an insert which is, apart from a single restriction site, identical to Ch2 (see below). In the case of cos 14 ligation artifacts can therefore be excluded. Also within the sequenced regions of cos 13, 52, and 55 (Fig. 2) no artifactual ligation has occurred, since there is no interruption of homology between the analysed $V_K$ regions. The appearance of small cosmid inserts points to the necessity to use carefully size fractionated DNA fragments treated with phosphatase and cosmid vector arms prepared as in ref. 36.



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

Sequence analysis of six $V_{KI}$ regions

The six $V_K$ regions located on the phage and cosmid clones of Fig. 1 were sequenced and compared to each other and to the sequence of the mouse $V_K$ probe which was used to isolate the phages Ch1 and Ch2 (Fig. 2).

The sequence of V1 turned out to be identical to the sequence of HK102 which had been isolated from the same phage library by D.L. Bentley and T.H. Rabbitts (37). Our sequence extends the published HK102 sequence by 381 bp in the 5' and 3' directions and is referred to as V1 in this paper.

The formal translation products of the six human $V_K$ sequences (Fig. 3) identify them as members of subgroup I. Only two sequences code for potentially functional $V_{KI}$ genes, V1 (HK102) and V13. V2, V14, V52, and V55 are pseudogenes (Fig. 3) which deviate in several features from functional genes. V2 and V14 lack an ATG codon in the expected position of the L region and carry two frameshift mutations relative to another ATG codon within L; in addition they have stop codons within FR3. V52 contains a frameshift mutation within FR3 and an altered VJ joining signal (9). V55 carries two frameshift mutations within FR1.

When the sequences of Fig. 2 were compared in the leader-intron-$V_K$ segment, the mouse gene was found to be 60 - 66 % homologous to the various human genes. The corresponding values for comparisons among the human genes are 69 - 99 %. Human $V_{KI}$ genes can therefore be about as homologous to mouse genes as to other human $V_{KI}$ genes. A similar observation has been made with human and mouse heavy chain genes (12).

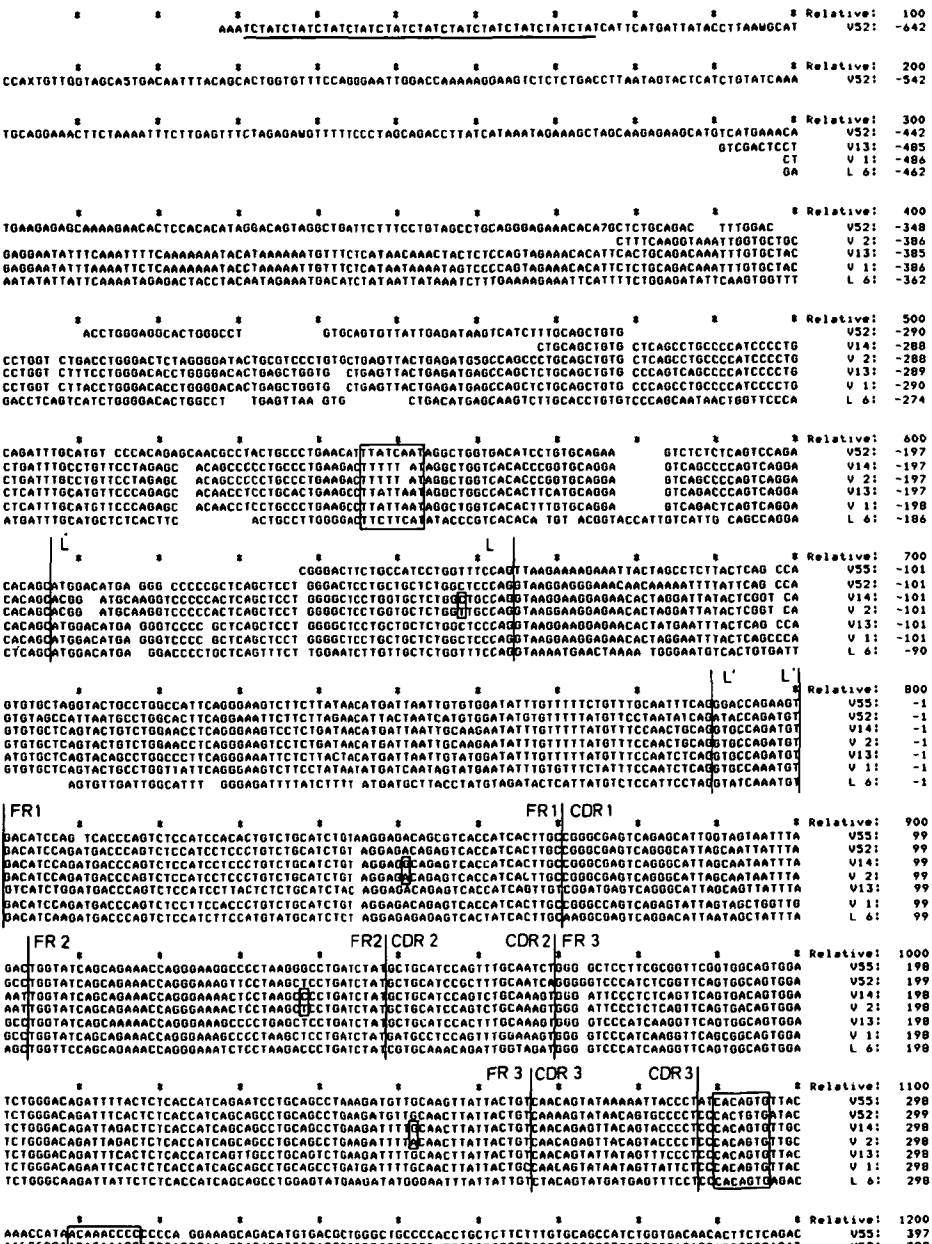
Alleles in a multigene family

The restriction maps of V2 and V14 are identical in a region of 6.6 kb apart from a single PvuII site in the 3' flank of V14. The sequences of V2 and V14, as far as they have been determined, differ only in 6 out of 703 positions (0.85 %). A low degree of divergence (3.5 %) is also present between V13 and the $V_{KI}$ gene Ve (15). V2 and V14 as well as V13 and Ve may be alleles or they may have originated from a common ancestor gene by a recent duplication event.

Restriction site differences together with appropriate pro-



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024


Figure 2. Sequences of six human $V_K$ sequences and comparison to the sequence of a mouse $V_K$ gene. The sequences of the human $V_K$ genes and pseudogenes V55, V52, V14, V2, V13, and V1 were determined by the chain termination method (31) following the strategies shown in Fig. 1. The mouse $V_K$ sequence L6 is taken from ref. 30. Sequences were aligned for maximum homology with the help of the sequence editor DNMAHO (Methods). The coding regions were identified by comparison with published amino acid sequences (41) and according to exon-intron splicing rules (42). It should be noted that the donor splice site of the pseudogene V55 differs from the consensus splice site. The TATA and VJ-joining signal sequences (42,9) are boxed. The six positions where V14 differs from V2 are also boxed. A simple sequence in the 5' flank of V52 is underlined. Two numbering systems are used: "relative" positions are identical for the seven aligned sequences and are the basis of sequence comparisons. The numbering of the individual sequences starts at the first nucleotide of FR1. The sequence of V1 is identical to the sequence of HK102 (37) from position -312 to +488 (our counting). V1 therefore extends the HK102 sequence (37) by 175 bp on the 5' and by 206 bp on the 3' direction. The following symbols represent positions which cannot be unambiguously assigned to a single nucleotide: W: A or G or T; S: A or C; X: A or C or T; 7: A or T. The subregion L' comprises codons -4 to -1.

bes can be used to distinguish between alleles and different members of a multigene family if the hybridization patterns of DNAs from a number of individuals are compared (15,38). The PvuII site in the 3' flank of V14 is not suitable for this test since a 3' flanking probe would detect too many crosshybridizing bands (35,15). In multigene families such as the $V_K$ gene family it is not possible to distinguish different members of the

Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

family from allelic variants if there are no diagnostic restriction sites available and the sequences are very similar.

Different sets of $V_{KI}$ gene sequences

With the six $V_{KI}$ sequences presented in this paper 16 different $V_{KI}$ gene sequences are known to date. In order to determine homologies and to find identical subregions the six $V_{KI}$ sequences V14, V2, V1, V13, V55, and V52 (this paper) were aligned with the published DNA sequences of Va,a',b,b',d,e (15), HK100, HK101 (37) and HK134, HK137 (38) with the help of the sequence editor DNMAHO (15). Using the computer program DNPERC (15) different subregions were compared in all possible combinations. This analysis clearly shows that the $V_{KI}$ subgroup can be further subdivided into sets of more closely related sequences. The coding regions of HK101, 134, 137, V1, 13, Va,a',b,b',d,e display homologies of 90-99.8 %. They form one set of related sequences which are also more similar in their 5' and 3' flanks than other $V_{KI}$ genes. Four other sets of $V_{KI}$ genes may be represented by the genes V14 and V2, V52, V55, and HK100, respectively. HK122 according to its partial sequence (39) seems to belong to the V14-V2 set. Homologies between the members of the altogether five different sets of $V_{KI}$ genes range from 77 to 85 %.

The $V_{KI}$ repertoire contains many pseudogenes

Two of the 16 $V_{KI}$ DNA sequences can be clearly regarded as allelic variants (HK101 and HK134; ref. 38). Among the 15 different

Figure 3. Comparison of amino acid sequences of $V_K$ genes and pseudogenes yields sets of identical subregions.
A) All published amino acid and nucleotide sequences of the human $V_{KI}$ subgroup are compared if their sequences were completely determined. The protein sequences are taken from ref. 41. The translated DNA sequences of HK100, 134, and 137 are from refs. 37 and 38, sequences of Va,a',b,b',d, and e are from ref. 15. The DNA sequences of V2, V14, V1, V52, V55, and V13 are shown in Fig. 2. The pseudogenes HK100, V2, V14, V52, V55, and Va were aligned for maximum homology in their DNA sequences and then formally translated: asterisks symbolize stop codons, points indicate frameshifts which have to be introduced to keep the amino acid sequences within the alignment. The numbering is chosen according to ref. 41.
B) Underneath the amino acid sequences the names of those sequences are boxed which are identical in the respective subregion. For example HK134 and Vb are identical in CDRI as are HK137 and V52.



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

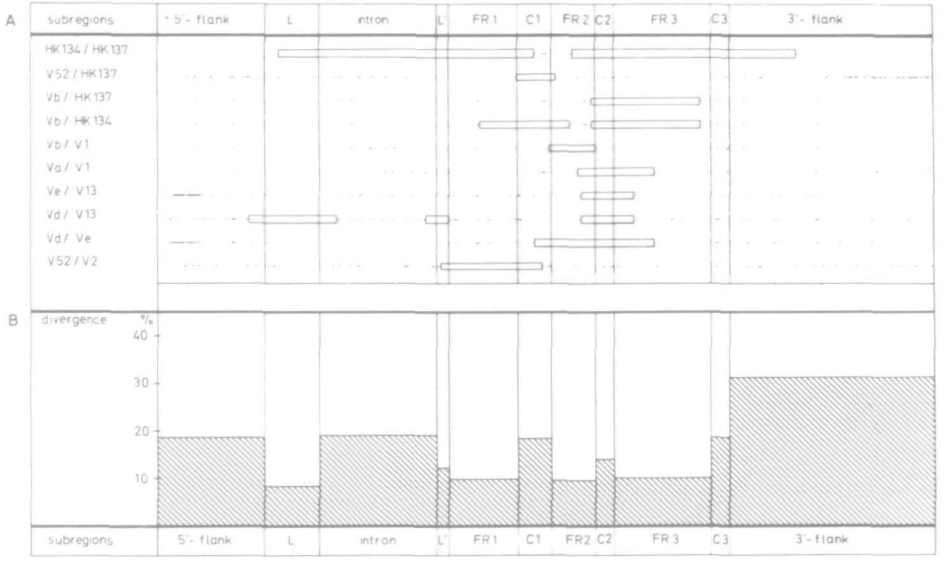



Figure 4. Divergence plots of $V_{KI}$ DNA sequences. 12 $V_{KI}$ sequences which appear not to be alleles (see text) were aligned for maximum homology (program DNMAHO, ref. 15) and all possible pairs were then screened for mismatches (a dot shows a difference of 1 bp, program DNCHOD).
A) In this panel only those comparisons are shown which result in identical sequences spanning at least one subregion. The identical sequences are indicated by open boxes. For the references to the various sequences see the legend of Fig. 3.
B) In this panel a histogram of mean divergences for the different subregions and flanking sequences is shown. The mean divergence in each subregion was determined by calculating the sum of divergences in all possible comparisons between the 12 sequences divided by the number of these comparisons. L' comprises codons -4 to -1; C is a further abbreviation of CDR.

$V_{KI}$ sequences are 6 pseudogenes (Va, HK100, V14, V2, V52, V55) and 9 potentially functional genes (Va', Vb, Vb', Vd, Ve, V1, V13, HK101, HK137). A large proportion of the $V_K$ repertoire may therefore consist of pseudogenes as was already suggested for the $V_H$ repertoire (40).

V2 and V14 differ in more features from potentially functional $V_K$ genes than for instance Va which is very similar to Va' (15). While Va is closely linked to a $V_{KI}$ gene cluster V2 and V14 may be more isolated genes. The degree of divergence of pseudogenes from functional genes may depend, in part, on



Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

their distance to other genes since the distance may influence the frequency of gene conversion-like events.

$V_{KI}$ sequences share identical subregions

It is interesting to discuss the data of all $V_{KI}$ amino acid and DNA sequences known up to now in relation to the minigene hypothesis of E.A. Kabat et al. (6). Fig. 3 shows a comparison of all complete $V_{KI}$ amino acid sequences derived from protein data (41) and from formal translation of DNA sequences (this paper, refs. 15, 37, 38). This compilation of $V_{KI}$ amino acid sequences (Fig. 3) was searched for subregions with identical sequences, thus extending the analysis of E.A. Kabat et al. (41) which includes also incomplete protein sequences but only two translated DNA sequences. Single V regions seem to be composed of subregions belonging to different sets. This may be demonstrated for one example: V1 is identical to the EU and DEN protein sequences in FR1; in FR2 it is identical to ROY, AU, KA, Vb, Vd, and Ve and in CDR2 it is identical to Va. The sequence of V1 may therefore be interpreted as the product of assembled minigenes. But the stretches of identical sequences do not coincide with subregions as can be most clearly seen in a comparison of twelve $V_{KI}$ DNA sequences (Fig. 4A). The patches of DNA sequences identical in different $V_{KI}$ genes reach beyond the borders of FRs and CDRs. Nevertheless, the subregions are still clearly distinguished by different degrees of divergence (Fig. 4B). It is noteworthy that the L region is about as conserved as the FRs.

Both structural features, the extension of the identical DNA patches and the different levels of divergence for FRs and CDRs can be explained by an intergenic exchange of information. The mechanism which is responsible for those structural features involves probably gene conversion-like events (11,14) and subsequent selection acting on sets of the $V_K$ sequences (15). These two processes may result in sequences which appear as if they had been composed from separately encoded minigenes (6,11). Data are accumulating which point to a role of gene conversion-like processes in both, the conservation of FRs and the generation of diversity in CDRs. But the extent of the contribution of such processes can be estimated only when a complete V gene locus is analysed and different haplotypes are compared.

Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

ACKNOWLEDGMENT

We are indebted to T. Maniatis for his phage library and to J. Collins in whose laboratory the cosmid libraries were prepared by W.L. and screened by M.P. We thank P.S. Neumaier for help in the computing and R. Gerl and E. Osterholzer for expert assistance. The work was supported by Bundesministerium für Forschung und Technologie and Fonds der Chemischen Industrie.

Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024

REFERENCES

1. Present address: Patentanwalts-Kanzlei Dr. Vossius, Siebertstrasse 4, 8000 München 80.
2. The following nomenclature is used in this and related papers. Abbreviations: V, variable, FR, framework, CDR, complementarity-determining regions; L, leader, J, joining, C, constant region of the immunoglobulin kappa light chain and the corresponding gene segments. Ch1 ff: clones with the lambda phage Charon 4A as vector isolated from the library of Lawn et al. (see below ref. 21). cos 10 to cos 49: clones with pHC79-2 cos (see below ref. 23) as vector isolated from cosmid library I. cos 50 to cos 99: clones with pHC79-2 cos/tk as vector isolated from cosmid library II. cos 100 ff: clones with pHC79 (see below ref. 25) as vector isolated from cosmid library III. DNA fragments, subclones and subsubclones are designated by numbers, e.g. 1-2 is a subclone from Ch1.
3. Cold Spring Harbor Symp. Quant. Biol. 42 (1977).
4. Wu, T.T. and Kabat, E.A. (1970) J. Exp. Med. 132, 211-250.
5. Davies, D.R. and Metzger, H. (1983) Ann. Rev. Immunol. 1, 87-117.
6. Kabat, E.A., Wu, T.T., and Bilofsky, H. (1978) Proc. Natl. Acad. Sci. USA 75, 2429-2433.
7. Tonegawa, S., Maxam, A.M., Tizard, R., Bernard, O., and Gilbert, W. (1978) Proc. Natl. Acad. Sci. USA 75, 1485-1489.
8. Seidman, J.G., Leder, A., Edgell, M.H., Pobky, F., Tilghman, S.M., Tiemeier, D.C., and Leder, P. (1978) Proc. Natl. Acad. Sci. USA 75, 3881-3885.
9. Tonegawa, S. (1983) Nature 302, 575-581.
10. Kabat, E.A., Wu, T.T., and Bilofsky, H. (1979) J. Exp. Med. 149, 1299-1313.
11. Wu, T.T. and Kabat, E.A. (1982) Proc. Natl. Acad. Sci. USA 79, 5031-5032.
12. Rechavi, G., Ram, D., Glazer, L., Zakut, R., and Givol, D. (1983) Proc. Natl. Acad. Sci. USA 80, 855-859.
13. Bernstein, K.E., Reddy, E.P., Alexander, C.B., and Mage, R.G. (1982) Nature 300, 74-76.
14. Baltimore, D. (1981) Cell 24, 592-594.
15. Pech, M., Jaenichen, H.R., Pohlenz, H.D., Neumaier, P.S., Klobeck, H.-G., and Zachau, H.G. (1984) J.Mol.Biol. in press.
16. Leder, P., Tiemeier, D., and Enquist, L. (1977) Science 196, 175-177.
17. Streeck, R.E. and Hobom, G. (1975) Eur. J. Biochem. 57, 595-606.

18. Boyer, H.W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.
19. Messing, J., Crea, R., and Seeburg, P.H. (1981) Nucl. Acids Res. 9, 309-321.
20. Messing, J. (1983) Meth. Enzym. 101, 20-78.
21. Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G., and Maniatis, T. (1978) Cell 15, 1157-1174.
22. Blin, N. and Stafford, D.W. (1976) Nucl. Acids Res. 3, 2303-2308.
23. Lindenmaier, W., Hauser, H., Greiser de Wilke, I., and Schütz, G. (1982) Nucl. Acids Res. 10, 1243-1256.
24. Lin, C.-P., Tucker, P.W., Mushinski, J.F., and Blattner, F. R. (1980) Science 209, 1348-1353.
25. Hohn, B. and Collins, J. (1980) Gene 11, 291-300.
26. Grosveld, F.G., Dahl, H.H.M., de Boer, E., and Flavell, R. A. (1981) Gene 13, 227-237.
27. Denhardt, D.T. (1966) Biochem. Biophys. Res. Commun. 23, 641-646.
28. Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K., and Efstratiadis, A. (1978) Cell 15, 687-701.
29. Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.
30. Pech, M., Höchtl, J., Schnell, H., and Zachau, H.G. (1981) Nature 291, 668-670.
31. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H., and Roe, B.A. (1980) J. Mol. Biol. 143, 161-178.
32. Anderson, S. (1981) Nucl. Acids Res. 9, 3015-3027.
33. Garoff, H. and Ansorge, W. (1981) Anal. Biochem. 115, 450-457.
34. Staden, R. (1980) Nucl. Acids Res. 8, 3673-3694.
35. Straubinger, B., Pech, M., Mühlebach, K., Jaenichen, H.R., Bauer, H.G., and Zachau, H.G. (1984) following paper.
36. Ish-Horowicz, D. and Burke, J.F. (1981) Nucl. Acids Res. 9, 2989-2998.
37. Bentley, D.L. and Rabbitts, T.H. (1980) Nature 288, 730-733.
38. Bentley, D.L. and Rabbitts, T.H. (1983) Cell 32, 181-189.
39. Bentley, D.L., Farrell, P.J., and Rabbitts, T.H. (1982) Nucl. Acids Res. 10, 1841-1856.
40. Cohen, J.B. and Givol, D. (1983) EMBO J. 2, 1795-1800.
41. Kabat, E.A., Wu, T.T., Bilofsky, H., Reid-Miller, M., and Perry, H. (1983) Sequences of Proteins of Immunological Interest, NIH Publication.
42. Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.

Downloaded from https://academic.oup.com/nar/article/12/13/5249/1034840 by guest on 19 April 2024