

GEN 02458

Localization, analysis and evolution of transposed human immunoglobulin V_κ genes

(Human-rodent cell hybrids; cosmids; restriction maps; ligation artifacts; orphon; recombinant DNA)

Erika Löttscher^a, Franz-Josef Zimmer^a, Thomas Klopstock^a, Karl-Heinz Grzeschik^b, Rita Jaenichen^a, Bernhard Straubinger^a and Hans G. Zachau^a

^a Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, 8000 München 2 (F.R.G.) and ^b Institut für Humangenetik der Universität Münster, 4400 Münster (F.R.G.)

Received 26 February 1988

Accepted 20 March 1988

Received by publisher 8 April 1988

SUMMARY

The localization of V_κ gene regions to chromosome 2, on which the κ locus is located, and to other chromosomes is described. The V_κ genes that have been transposed to other chromosomes are called orphans. The finding of two new V_κ genes on chromosome 22 is reported. A V_κ II gene of this region and two V_κ I genes of the Chr1 and the cos 118 regions were sequenced. The two V_κ I orphon sequences and two others that had been determined previously were 97.5% identical, indicating that they may have evolved from a common ancestor by amplification. A model of the evolution of the human V_κ orphans is discussed.

INTRODUCTION

The light chains of the κ type of the human immunoglobulins are encoded by numerous V_κ gene segments and by one J_κ-C_κ gene segment. Functional κ genes are formed by recombination of one V_κ gene segment with J_κC_κ (for reviews, see

Tonegawa, 1983; Honjo and Habu, 1985). A V_κ gene of subgroup I and the C_κ gene segment has been localized to chromosome 2 (Malcolm et al., 1982; McBride et al., 1982). We have established a physical map of the human κ locus comprising about 80 V_κ genes, in addition to the J_κC_κ gene segment (Lorenz et al., 1987). In the course of our cosmid

Correspondence to: Dr. H.G. Zachau, Institut für Physiologische Chemie, Goethestrasse 33, 8000 München 2 (F.R.G.) Tel. (089)5996429.

Present addresses: (E.L.) Howard Hughes Medical Institute, Department of Medicine, University of California, San Francisco, CA 94143 (U.S.A.) Tel. (415)476-1860; (K.-H.G.) Institut für Humangenetik, Bahnhofstrasse 7a, 3550 Marburg (F.R.G.) Tel. (06421)284080; (B.S.) Department of Plant Pathology, Cornell University, Ithaca, NY 14853 (U.S.A.) Tel. (607)255-3200.

Abbreviations: aa, amino acid(s); bp, base pair(s); Chr1, Chr22, V_κ gene-containing regions of chromosomes 1 and 22; FR, CDR, framework and complementary determining regions; kb, kilobase(s) or 1000 bp; L, L', parts of a leader gene segment; m219-1, the first subclone of the cosmid clone cos 219; orphon, V_κ gene outside the κ locus on chromosome 2p12; SSC, 0.15 M NaCl, 0.015 M Na₃·citrate, pH 7.6; V, J, C, variable, joining and constant gene segments; V_κ I to V_κ IV, variable gene segments of immunoglobulin light chains of the κ type belonging to subgroups I to IV; for reasons of simplicity V_κ gene segments are generally called V_κ genes.

cloning experiments we encountered a number of V_{κ} genes which lie outside the κ locus and which we call orphans, in analogy to the histone and ribosomal RNA genes outside of the respective gene clusters (Childs et al., 1981). Some orphan V_{κ} genes have been localized on chromosomes 1 and 22 (Lötscher et al., 1986) and on other chromosomes (Straubinger et al., 1988a). A clustering of putative autonomously replicating sequences has been detected in the 5' flanks of some V_{κ} I orphans indicating that they may have been formed in evolution by over-replication and transposition events (Lötscher et al., 1988).

The finding that V_{κ} genes are located not only within the κ locus but also outside it made it necessary to determine the chromosomal localization of all known human V_{κ} gene regions. The pertinent experiments are described in the present paper together with structural data on the orphan V_{κ} genes. The chromosomal location studies also turned out to be a reliable way to detect ligation artifacts in cosmids which had arisen, as rare events, in the preparation of a cosmid library by ligation of V_{κ} gene-containing fragments with DNA fragments derived from other chromosomes. The structural similarities between the transposed V_{κ} genes of chromosomes 1 and 22 and of the so-called Z family (Straubinger et al., 1988) led to speculations on their evolutionary origin.

MATERIALS AND METHODS

The materials and methods used in the present study were the same as in Straubinger et al. (1988a). In addition, some of the hybridization experiments were carried out in phosphate buffer as described by Church and Gilbert (1984). The mouse-human and hamster-human cell hybrids which were used for the chromosomal location analyses were the same as in Lötscher et al. (1986). The cosmid clones cos 206 and 218 are derived from the cosmid library III (Pohlenz et al., 1987).

RESULTS AND DISCUSSION

(a) Chromosomal localization of V_{κ} gene regions

In the course of the elucidation of the human κ locus, we analyzed numerous V_{κ} gene-containing

cosmids isolated from genomic DNA libraries. The analyses revealed that some of the V_{κ} gene regions occur outside the κ locus. To continue the structural work on the human κ locus successfully, it was necessary to check systematically the chromosomal localization of V_{κ} regions. The method for the chromosomal localization experiments was the same as used previously (Lötscher et al., 1986). Briefly, specific subclones prepared from the various regions were hybridized to a panel of digested DNAs from human-rodent cell hybrids containing different sets of human chromosomes. The segregation analysis of the fragments and human chromosomes in the panel allows one to assign the regions detected by the probes to specific chromosomes. The essential part of the results is shown in Fig. 1. A full account of the localization experiments is given in Lötscher (1988).

As expected, all of the regions which contained several V_{κ} genes mapped to chromosome 2. This is shown in Fig. 1 for the A regions (Straubinger et al., 1987) and the O and W regions (Pohlenz et al., 1987). The assignment of the L regions to chromosome 2 has been described previously (Lötscher et al., 1986).

(b) Identification of cosmid ligation artifacts by chromosomal localization experiments

Even if all steps in the preparation of a cosmid library are carefully controlled, a small amount of ligation artifacts is encountered, which we estimate to be about 3% (Pohlenz et al., 1987). Chromosomal localization experiments allowed us now to identify two such artifacts.

The first example is the cosmid clone cos 219 (Fig. 2). The subclone m219-1 was clearly localized to chromosome 5 (Fig. 1) but the region upstream from the V_{κ} II gene hybridized strongly to the subclone m654-1 of the Wb region (Pohlenz et al., 1987) which is part of chromosome 2 (see above). The sequence of the V_{κ} II gene is identical to that of W1 except for a T instead of a C in the L-L' intron (T. Klopstock, unpublished). The left part of cos 219 is therefore probably derived from an allele of the Wb region and the right part from chromosome 5. No vector sequences were found by hybridization experiments at the point of ligation (indicated in Fig. 2).

Another example of a ligation artifact is the V_{κ} I gene-containing clone cos 118 which was thought to

Human chromosomes																								Genomic regions				
Hybrids	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	A region	O region	W region	cos 219
RAG Anly 1									▲					▲										▲				
A9 SU 1-2	■	▲		▲											▲													
V79 Ly 3-2	■				■																							
V79 Ly 3-3										■																		
RAG GM 610 5-23															■									▲				
A9 Call 1-9-9	■	▲		▲											▲													
A9 Call 1-13																												
RAG RU 1a-5-11									▲																			
A9 JT 2-21-14	■	▲		▲											■									▲				
RAG GM 97 8-13	■									■																		

Fig. 1. Chromosomal location of V_{κ} gene-containing regions. The chromosomal locations of the A, O and W regions and of cos 219 (see section a) were carried out using the self-ligated inserts of the clones m163-1 (Straubinger et al., 1987), m143-1 and m167-1 (Pohlentz et al., 1987) and m219-1 (Fig. 2), respectively. The hybrid cell lines are the same as those used in Lötscher et al. (1986). Filled spaces indicate the presence of the chromosome in the respective cell line; half-filled spaces mean that part of the chromosome is present, and empty spaces indicate that the chromosome was not detected by biochemical and cytogenetic analyses.

be located on chromosome 15 on the basis of hybridization experiments with a subclone from the left side of the cosmid clone (Lötscher et al., 1986). It was then shown, by hybridization with parts of the vector pHc79 and by sequencing, that a fragment near the break-off of homology to the Chr1 map (Fig. 1 in Lötscher et al., 1986) is derived from the vector. It was not possible to prepare a single copy subclone from the right side of the cosmid clone which would permit the chromosomal localization of the V_{κ} I gene region of cos 118. However, according to its map (Lötscher et al., 1988) and the sequence of the gene

(see below) it is closely related to both Chr1 and the Z family (Straubinger et al., 1988). Apparently an orphon V_{κ} I gene region had been artifactually ligated to a vector fragment and to a fragment derived from chromosome 15. The experiments which led to the clarification of this situation have been documented (Lötscher, 1988).

The lesson learned from these two artifacts is that chromosomal localization analyses should be done for all cosmid clones which are isolated only once from a library and which contain long regions free of genes.

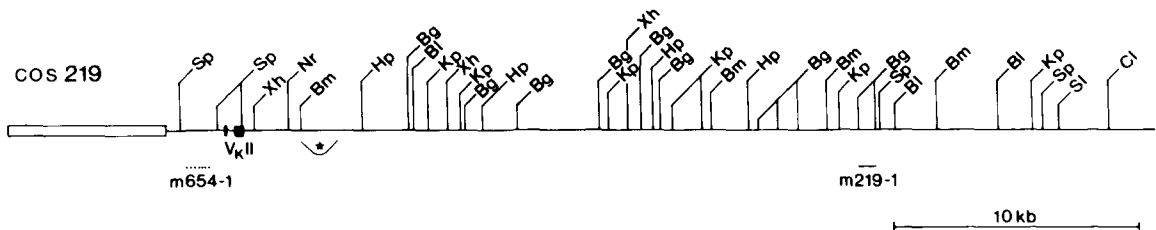


Fig. 2. Restriction map of the cos 219 region. The restriction map shows the product of the ligation of two different genomic fragments. The site of ligation is within the bracket indicated by an asterisk. The subclone m219-1 marked by a line was localized to chromosome 5 (Fig. 1). The subclone m654-1 derived from the Wb regions on chromosome 2 (Pohlentz et al., 1987) hybridizes to the fragment of cos 219 indicated by a short dotted line. The V_{κ} II gene sequence and the restriction map around the gene are identical to those of the gene W1 (Pohlentz et al., 1987). The abbreviations of the restriction nucleases are the same as in Fig. 3.

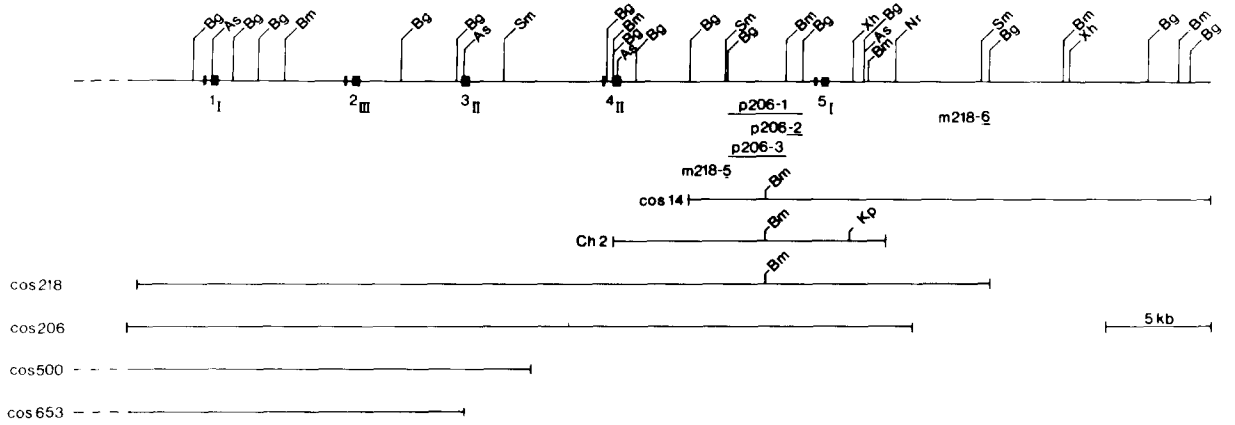


Fig. 3. Restriction map of part of the genomic region Chr22. The restriction map contains an extension to the 3' direction compared to the one by Löttscher et al. (1986). The V_x genes and leaders are indicated by filled boxes; the suffixes at the gene numbers specify the subgroups. Horizontal bars show the lengths of the cosmids. The presence of additional restriction sites which occur only in the respective cosmids are indicated. The designations of the clones cos 500 and cos 653 are the same as in Bauer et al. (1986). The subclones are shown as short horizontal lines. The fragments 206-1 to 3 were subcloned into the vector ptk (Holst et al., 1988). Restriction nucleases are abbreviated: As, *Asp*718; Bg, *Bgl*II; Bl, *Bgl*I; Bm, *Bam*HI; Cl, *Cla*I; Dr, *Dra*I; Ec, *Eco*RI; Hp, *Hpa*I; HIII, *Hind*III; Kp, *Kpn*I; Nr, *Nru*I; Ps, *Pst*I; Pv, *Pvu*I; PII, *Pvu*II; Sl, *Sal*I; Sm, *Sma*I; Sp, *Sph*I; Xh, *Xho*I.

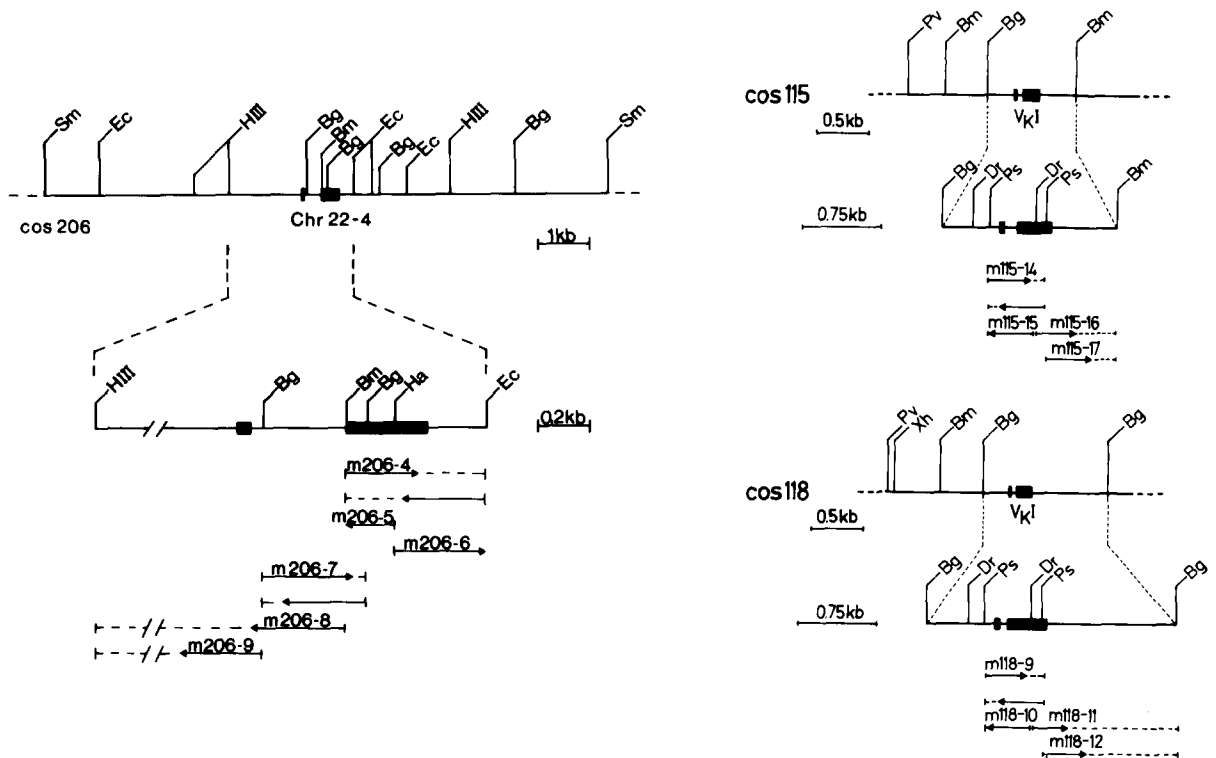


Fig. 4. Restriction maps and sequencing strategies of V_x genes of the Chr1 (cos115) and Chr22 (cos206) genomic regions and of cos 118. Gene segments are shown as filled boxes. Extent and direction of sequencing are symbolized by arrows. The horizontal dashed lines indicate the lengths of the fragments cloned into M13 vectors; arrows indicate the extent of sequencing. The abbreviations of the restriction nucleases are the same as in Fig. 3.

(c) The Chr22 region contains five V_{κ} genes

Most of the V_{κ} regions located outside the κ locus contain a single V_{κ} gene of the subgroup I (Lötscher et al., 1986; Straubinger et al., 1988a). However, a cluster of three V_{κ} genes could be shown to be part of a region which has been transposed to chromosome 22 (Lötscher et al., 1986). The restriction map of the Chr22 region could be extended now in the 3' direction by the two cosmid clones cos 206 and cos 218 which were isolated from genomic DNA libraries with the help of two V_{κ} II-specific probes (m607V-3; Klobeck et al., 1984, and m106-1, a 448-bp *Pst*I clone comprising part of the V_{κ} '' gene; Pech et al., 1985). In addition, the λ Charon phage Ch2 and the cosmid clone cos 14 which had been isolated previously (Jaenichen et al., 1984) overlap with the extended restriction map of the Chr22 region. The extending phage and cosmid clones were assigned to chromosome 22 on the basis of restriction map homologies and hybridization experiments. The subclone m218-6 (Fig. 3) which was prepared for further chromosomal walking in the 3' direction contained low repetitive sequences and therefore could not be used.

Ligation artifacts are rather unlikely in the Chr22 region since most of it is derived from several overlapping clones (Fig. 3, and Lötscher et al., 1986). In addition, a subclone prepared from the region upstream from the gene Chr22-2 (m120-12 in Bauer et al., 1986) was found to hybridize to the DNA of chromosome 22-containing hybrid cell lines (not shown). That the subclone hybridized also to the chromosome 2-containing cell lines is explained by its proximity to a V_{κ} III gene and the general homology of flanking sequences of V_{κ} genes of certain subgroups.

The new 3' part of the Chr22 region turned out to contain two additional V_{κ} genes. The sequence of the V_{κ} II gene Chr22-4 is reported below. The sequence of an allele of the V_{κ} I gene Chr22-5 had been determined previously in the phage clone Ch2 (V2 in Jaenichen et al., 1984). It is interesting that the Chr22-5 gene and its 5' and 3' flanks are highly homologous to the V_{κ} I gene regions of other orphans (Lötscher et al., 1988).

(d) Nucleotide sequences of three transposed V_{κ} genes

To determine the degree of homology between transposed genes and genes located in the κ locus, the gene Chr22-4 of the Chr22 region, as well as the genes of the Chr1 and cos 118 regions, were sequenced. The sequencing strategies are shown in Fig. 4. The sequence of the gene Chr22-4 (Fig. 5) is very similar to the V_{κ} II genes of the κ locus. Only one of the 57 invariant aa (Kabat et al., 1987) is changed: at aa position 7 of FR1 the serine is replaced by a threonine. The regulatory sequences and recognition signals correspond to the respective consensus sequences. However, at position 81 of FR3, the stop codon TAG was found. Besides the chromosomal localization of the gene on chromosome 22 the occurrence of the stop codon is the only indication which points to a non-functional V_{κ} gene.

The V_{κ} I genes of both the Chr1 and cos 118 regions contain a mutation in the start codon (Fig. 6). It should be noted, however, that the third codon is a start or methionine codon. The regulatory sequences of the two genes correspond to the respective consensus sequences. Three and four of the 29 invariant amino acids of subgroup I (Kabat et al., 1987) are changed in the genes of the Chr1 and cos 118 regions, respectively. The latter gene also contains a TAG stop codon in FR3.

(e) Comparison of the sequences of four transposed V_{κ} I genes

A characteristic feature of all transposed V_{κ} regions analyzed so far is the presence of V_{κ} I gene regions with very similar restriction maps. To determine how far the homology extends to the sequence level the sequences of the V_{κ} I genes of the Chr1, the cos 118, and the Z1 regions were compared to the sequence of the V2 gene which corresponds to Chr22-5 (Fig. 6). In addition to the coding sequences, 112 bp of the 5' flank and 175 bp of the 3' flank were included in the comparison. The overall identity proved to be 97.5%. This number is in the same range as is found for allelic or duplicated V_{κ} genes within the κ locus. In contrast, the average homology between randomly chosen V_{κ} I genes is about 90%. In addition, four identical mutations in comparison to a potentially functional V_{κ} I gene



Fig. 5. Nucleotide sequence and deduced amino acid sequence of the V_κ II gene Chr22-4. The sequence was determined following the strategy shown in Fig. 4. Invariant amino acids of subgroup II (Kabat et al., 1987) are underlined. A horizontal line above an aa symbol indicates that an invariant amino acid has been replaced by another amino acid. A conserved pentekaidecanucleotide, the decanucleotide promoter (Falkner and Zachau, 1984), a TATA box and the hepta- and nonanucleotide boxes are underlined. The asterisk marks a stop codon. The numbering and the division into L, L', FRI etc. is as in Kabat et al. (1987).

have been found in the V_κ I orphans: the start codon is changed in each case to ACG or ACA and 3 invariant aa have been replaced by the same substitute amino acid (Fig. 6). The high sequence homology as well as the identical mutations in the transposed sequences led to the hypothesis that the human V_κ orphans stem from a common ancestor.

(f) Concluding remarks on the evolution of the orphon V_κ genes

A hypothetical scheme of the evolution of the orphon V_κ genes is outlined in Fig. 7. We assume that the five genes of Chr22 have been transposed in a single event to chromosome 22 since it is unlikely

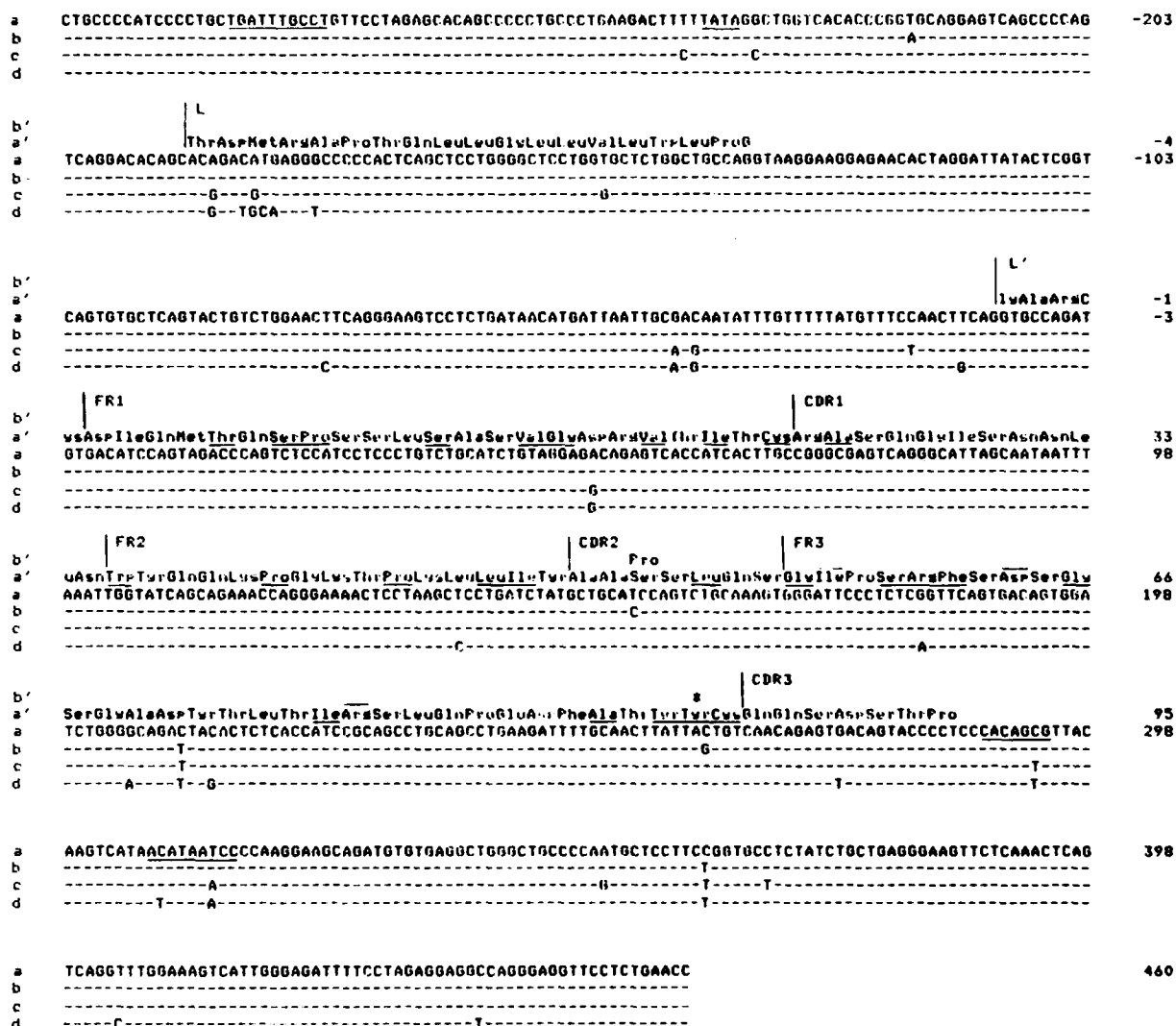


Fig. 6. Sequence comparison of four orphon V_{κ} I genes. The sequences of the orphon V_{κ} I genes of the genomic regions of Chr1 (a) and cos 118 (b) were determined following the strategies shown in Fig. 4. The respective formal translation products a' and b' are also shown. The gene of the ZI region (c) and the allele V2 of the Chr22-5 gene (d) are taken from Straubinger et al. (1988a) and Jaenichen et al. (1984), respectively. Sequences were aligned for maximum homology with the help of sequence editor DNMAHO (Pech et al., 1984). The numbering, partitioning by vertical lines, the asterisk and lines above and below the symbols are as in Fig. 5.

that several transposition events have occurred independently to the same region of one chromosome. Whether the Chr22 transposition event has been duplicative or not is not clear. However, we have not so far found an apparent ancestral region within the κ locus, either by hybridization experiments or comparison of restriction maps.

The conserved V_{κ} I gene regions of the other orphans have been transposed to the other chromosomes from either the Chr22 region in a duplicative event or directly from the κ locus in a duplicative or

non-duplicative event. It is possible that the Chr1 and Z families have a common intermediate. In evolution there have probably been several amplification events involving the conserved V_{κ} I gene regions, since their co-amplified surroundings are different in ZI-ZIV (Straubinger et al., 1988), and in Chr1 also there are at least two types of constituent cosmids (Pohlenz, 1986). A clustering of putative replicating and/or recombinogenic elements upstream from the V_{κ} I orphans, described in an accompanying paper (Löttscher et al., 1988), may have

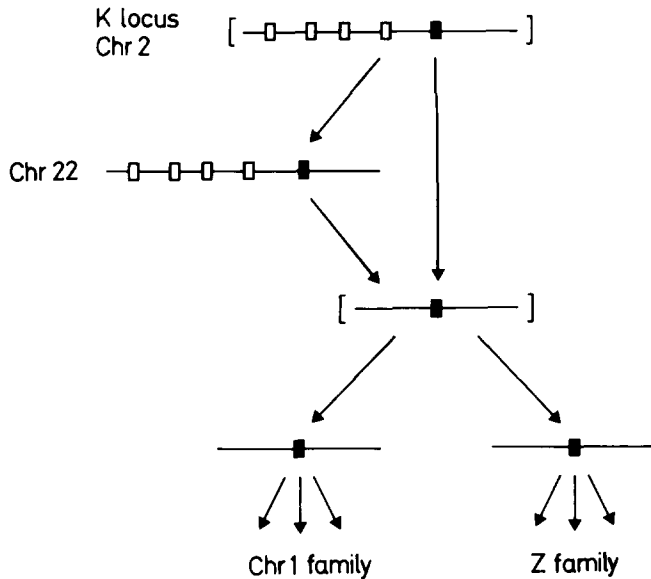


Fig. 7. Hypothetical model of the evolution of the human orpho V_{κ} genes (section f of the text). Open boxes indicate the Chr22-1 to 4 genes, the filled boxes the conserved $V_{\kappa} I$ genes of Chr22, Chr1, and the Z regions. The regions in square brackets indicate the hypothetical ancestor on chromosome 2 and a possible intermediate. The gene-containing parts of the Chr22, Chr1 and Z regions are compiled in Fig. 1 of Löttscher et al. (1988). The maps of Chr22 and Chr1 are described in Löttscher et al. (1986), those of the Z regions in Straubinger et al. (1988b).

been the cause of their over-replication and eventual transposition during evolution.

ACKNOWLEDGEMENTS

The work was supported by Bundesministerium für Forschung und Technologie and Fonds der Chemischen Industrie. F.-J.Z. is the holder of an A. Butenandt fellowship. K.-H.G. acknowledges the technical assistance of G. Schwartenbeck.

REFERENCES

Bauer, H.G., Pech, M. and Zachau, H.G.: Characterization of a group of transposed human V_{κ} genes. *Biol. Chem.* 367 (1986) 751–755.
 Church, G.M. and Gilbert, W.: Genomic sequencing. *Proc. Natl. Acad. Sci. USA* 81 (1984) 1991–1995.

Falkner, F.G. and Zachau, H.G.: Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. *Nature* 310 (1984) 71–74.
 Holst, A., Müller, F., Zastrow, G., Zentgraf, H., Schwender, S., Dinkl, E. and Grummt, F.: Murine genomic DNA sequences replicating autonomously in mouse L cells. *Cell* 52 (1988) 355–365.
 Honjo, T. and Habu, S.: Origin of immune diversity: genetic variation and selection. *Annu. Rev. Biochem.* 54 (1985) 803–830.
 Jaenichen, H.R., Pech, M., Lindenmaier, W., Wildgruber, N. and Zachau, H.G.: Composite human V_{κ} genes and a model of their evolution. *Nucleic Acids Res.* 12 (1984) 5249–5263.
 Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H. and Gottesman, K.S.: Sequences of proteins of immunological interest. National Institutes of Health, Bethesda, MD, 1987.
 Klobeck, H.-G., Solomon, A. and Zachau, H.G.: Contribution of human $V_{\kappa} II$ germ-line genes to light-chain diversity. *Nature* 309 (1984) 73–76.
 Lorenz, W., Straubinger, B. and Zachau, H.G.: Physical map of the human immunoglobulin κ locus and its implications for the mechanisms of V_{κ} - J_{κ} rearrangement. *Nucleic Acids Res.* 15 (1987) 9667–9676.
 Löttscher, E., Grzeschik, K.-H., Bauer, H.G., Pohlenz, H.-D., Straubinger, B. and Zachau, H.G.: Dispersed human immunoglobulin κ light-chain genes. *Nature* 320 (1986) 456–458.
 Löttscher, E., Siwka, W., Zimmer, F.-J., Grummt, F. and Zachau, H.G.: Transposed human immunoglobulin V_{κ} genes carry clusters of putative autonomously replicating sequence elements. *Gene* 69 (1988) 225–236.
 Löttscher, E.: Menschliche Immunglobulingene vom κ -Typ: Variable Gensegmente innerhalb und ausserhalb des κ -Locus. Ph. D. Thesis, Fakultät für Biologie der Universität München, München, 1988.
 Malcolm, S., Barton, P., Murphy, C., Ferguson-Smith, M.A., Bentley, D.L. and Rabbitts, T.H.: Localization of human immunoglobulin κ light chain variable region genes to the short arm of chromosome 2 by in situ hybridization. *Proc. Natl. Acad. Sci. USA* 79 (1982) 4957–4961.
 McBride, O.W., Hieter, P.A., Hollis, G.F., Swan, D., Otey, M.C. and Leder, P.: Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes. *J. Exp. Med.* 155 (1982) 1480–1490.
 Pech, M., Jaenichen, H.R., Pohlenz, H.-D., Neumaier, P.S., Klobeck, H.-G. and Zachau, H.G.: Organization and evolution of a gene cluster for human immunoglobulin variable regions of the kappa type. *J. Mol. Biol.* 176 (1984) 189–204.
 Pech, M., Smola, H., Pohlenz, H.-D., Straubinger, B., Gerl, R. and Zachau, H.G.: A large section of the gene locus encoding human immunoglobulin variable regions of the kappa type is duplicated. *J. Mol. Biol.* 183 (1985) 291–299.
 Pohlenz, H.-D.: Organisation menschlicher Immunglobulingene. Isolierung und Analyse von Regionen des menschlichen V_{κ} -Locus durch Cosmidklonierung. Ph. D. Thesis, Fakultät für Chemie und Pharmazie der Universität München, München, 1986.

- Pohlenz, H.-D., Straubinger, B., Thiebe, R., Pech, M., Zimmer, F.-J. and Zachau, H.G.: The human V_{κ} locus: characterization of extended immunoglobulin gene regions by cosmid cloning. *J. Mol. Biol.* 193 (1987) 241–253.
- Straubinger, B., Huber, E., Lorenz, W., Osterholzer, E., Pargent, W., Pech, M., Pohlenz, H.-D., Zimmer, F.-J. and Zachau, H.G.: The human V_{κ} locus: characterization of a duplicated region encoding 28 different immunoglobulin genes. *J. Mol. Biol.* 199 (1988a) 23–34.
- Straubinger, B., Thiebe, R., Pech, M. and Zachau, H.G.: The Z family, a group of transposed human immunoglobulin V_{κ} genes. *Gene* 69 (1988b) 209–214.
- Tonegawa, S.: Somatic generation of antibody diversity. *Nature* 302 (1983) 575–581.

Communicated by Z. Hradečná.