Multiple components in restriction enzyme digests of mammalian (insectivore), avian and reptilian genomic DNA hybridize with murine immunoglobulin V_H probes

G.W.Litman, L.Berger and C.L.Jahn*

Sloan-Kettering Institute for Cancer Research, Walker Laboratory, Rye, NY 10580, and *Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309, USA

Received 12 March 1982; Revised and Accepted 6 May 1982

ABSTRACT

High molecular weight genomic DNAs isolated from an insectivore, Tupaia, and a representative reptilian, Caiman, and avian, Gallus, were digested with restriction endonucleases transferred to nitrocellulose and hybridized with nick-translated probes of murine V_H genes. The derivations of the probes designated S107V (1) and μ 107V (2,3) have been described previously. Under conditions of reduced stringency, multiple hybridizing components were observed with Tupaia and Caiman; only μ 107V exhibited significant hybridization with the separated fragments of Gallus DNA. The nick-translated S107V probe was digested with Fnu4H1 and subinserts corresponding to the 5' and 3' regions both detected multiple hybridizing components in Tupaia and Caiman DNA. A 5' probe lacking the leader sequence identified the same components as the intact 5' probe, suggesting that V_H coding regions were being detected. The results indicate that species as phylogenetically distant as the reptilians may possess multiple genetic components which exhibit significant homology with murine immunoglobulin V_H regions.

INTRODUCTION

Recent studies have indicated that a number of unique (genetic) mechanisms account for variation in the amino terminal portions of immunoglobulin heavy and light chains which contain the antibody combining site (1,4-6). The variable region of the heavy chain reflects the contributions of V_H, D_H and J_H gene segments which in the germline appear to be arrayed in tandem clusters 5' to the constant region genes (7-10). Estimates of 40-50 and 80 individual V_H genes in mouse (3) and humans (3,7), respectively, imply a significant germline contribution to antibody diversity. Somatic reorganization of the V_H and the 3' situated D_H and J_H segments effects combinatorial amplification (1,6) of the germline sequences which are altered further by somatic mutation (11,12). As a first step towards extending the characterization of immunoglobulin V_H genes below the level of the higher mammals, we have sought to identify components in restriction enzyme digests of genomic DNA isolated from a primitive insectivore (Tupaia) and representative avian (Gallus) and reptilian (Caiman)

species which cross-hybridize with murine V_H probes. We conclude that multiple sequences related to V_H genes are present in species as phylogenetically distant as the reptilians.

MATERIALS AND METHODS

DNA isolation

Animals were euthanized, the entire stomach and intestinal tract were removed and either the intact animal or the liver was frozen in liquid N_2 prior to DNA isolation using a modification of a previously described method (13). The modified procedure involves initial phenol extraction, followed by an ether extraction which is carried out prior to digestion with proteinase K for 5 hr at 37°C in 0.5% SDS. DNA was judged to be >100 kb when examined by agarose electrophoresis. Values for the DNA content/cell were taken from (14) or were determined experimentally (for Tupaia) by microdensitometric analysis of fixed, Feulgen stained cells.

Southern blotting

10 μ g of DNA was digested to completion with Eco RI, Bam HI or Sst I and subjected to electrophoresis in 0.8% agarose relative to Hind III digested λ phage and Hae III digested ϕ X174 standards. The vendor's recommendations with regard to reaction conditions were followed. The efficiency of digestion was monitored by direct inspection and photography of the ethidium bromide-stained gels. The use of additional enzyme and extended reaction times failed to affect the final hybridization patterns. Similarly, increased DNA loads did not reveal the presence of additional hybridizing components and decreased loads did not result in resolution of any closely spaced components. The digests were transferred to nitrocellulose (15) using 10 X SSC (1 X SSC = 0.15 M NaCl, 0.015 M Na citrate) and in each experiment the efficiency of transfer was verified by reanalysis of the blotted gel. The prehybridization and hybridization conditions (6 X SSC, 0.5% SDS, 1 X Denhardt's solution [16]), including the use of competing DNAs, RNAs, and polynucleotides in the latter, were as described (17). Hybridization was carried out at 65°C for 20 hr. Depression of hybridization temperature was not found to facilitate identification of weak hybridizing components vs background. The stringency of washing was varied from 3 X SSC - 0.1 X SSC, 0.1% SDS at 52°C and 1 X SSC, 0.1% SDS was found to be optimal for visualizing the multiple hybridizing components in the lower vertebrates. In some instances, nitrocellulose filters were rehybridized following removal of bound probe (experimentally verified) by two successive 20 min washes in H_2O at 65°C. Radioautographs were exposed for 3-10 days at

-70°C with Kodak XAR-5 film and a Cronex intensifying screen.

Probes of immunoglobulin V_H regions

The pS107V-1 (1) and p μ 107 (2,3) plasmids generously supplied by Drs. L. Hood and T. Rabbits, respectively, were derived from cDNAs and used to transform *E. coli* LE392 or HB101 as described (18). Both plasmids were isolated preparatively by CsCl centrifugation followed by gel filtration on Sephacryl S-300. The inserts derived by Pst I digestion of pS107V-1 is referred to as S107V and the insert derived by Eco RI/Bgl II digestion of p μ 107 (2) is referred to as μ 107V. The V_H coding region of S107V and μ 107V specify antibody to phosphorylcholine and a sheep red blood cell antigen, respectively. Both were recovered from preparative 5% polyacrylamide gels by electroelution. Isolated probes were nick-translated as described (19) employing α -³²P dTTP or deoxytrinucleotide phosphate pairs, e.g., dTTP plus dCTP (New England Nuclear). Subinserts of S107V either were derived by complete cleavage with AluI or Fnu4HI (New England Biolabs) followed by preparative electrophoresis in 7.5% polyacrylamide and nick-translated directly employing reduced concentrations of DNase I and shortened reaction times or were isolated from the nick-translated parent probe by restriction enzyme digestion. The labeled digests were ethanol-precipitated in the presence of carrier yeast tRNA at -70°C and subjected to polyacrylamide gel electrophoresis. The distribution of the individual components was visualized by a 5-10 min direct radioautograph of the gel slab. Subinserts were recovered by gentle rocking of the polyacrylamide slices in 2.0 ml of 200 mM NaCl, 10 mM Tris-HCl, 10 mM EDTA pH 7.4, which subsequently was heated (with the appropriate components of the hybridization mixture) to 100°C for 10 min and quick-chilled prior to addition to the remainder of the hybridization mixture. Restriction mapping confirmed the identity of the subinserts. All of the results discussed below were repeated on at least three separate occasions with different preparations of DNA and labeled probe.

RESULTS

The S107V-1 probe encodes nearly an entire variable region, as well as a 5' portion not represented in the mature protein chain (1) (Fig. 1). The relative boundaries of the hypervariable regions, as well as the restriction enzyme cleavage sites which define subinserts of the probe also are indicated. The hybridization patterns of nick-translated S107V with Sst I-digested genomic DNA isolated from four different species are shown in Fig. 2. The multiplicity of hybridizing components with BALB/c DNA (homologous species) confirms earlier

Fig. 1 V_H probe and subinserts representing the S107 phosphorylcholine binding heavy chain variable region (1). Positions in the amino acid sequence are presented in (a) and regions corresponding to untranslated (UT), leader (L), variable (V), diversity (D) and abbreviated joining (J') gene segments are indicated. The amino acids representing the termination point of the regions are numbered. The boundaries of the hypervariable regions (hv) are indicated in (b). The S107V-1 probe (referred to as S107V) and location of the cleavage sites for Fnu4H1 (F) and AluI (A) are indicated in (c). The boundaries and designations for the subinserts yielded by cleavage with these enzymes are indicated in (d) and (e). Approximate scale drawing (nt = nucleotide); amino terminal and 5' ends to the left.

reports (1,3,6-8,11,12,20,21) and at least in part reflects, under conditions of relaxed stringency, intrasubgroup homology (3,7). Fewer weaker hybridizing components were noted in genomic digests of Tupaia, relatively little hybridization (not visible in photographs) was observed with Gallus and a number of components varying in hybridization intensity were observed with Caiman. The intensity of hybridization of the stronger bands in Caiman was equivalent to the weaker bands in Mus. The number of hybridizing components in Sst 1 vs Bam H1 digests were equivalent for both Tupaia and Caiman (not illustrated).

A single recognition site for Fnu4H1 in S107V defines probes corresponding to 5' and 3' coding regions (Fig. 1). The hybridization patterns with the separated, nick-translated subinserts are shown in Fig. 3. The same components were detected in Tupaia and Caiman DNA employing S107V and S107V-F1, which

Fig. 2 Cross-hybridization of S107V (Fig. 1) with Sst I-digested DNA isolated from (a + b) *Mus musculus* (BALB/c mouse), (c) *Tupaia belangerii* (tree shrew), (d) *Gallus domesticus* (chicken), and (e) *Caiman crocodylus* (Caiman). Filter hybridization conditions are described in the text; the specific activity of the probe was 95 cpm/pg. When the S107V insert was nick-translated to a specific activity of 250 cpm/pg, the same components in (c) and (e) were detected and very weak hybridization (corresponding to that observed with μ 107V, Fig. 4) was detected in (d). In order to better distinguish hybridizing components in *Mus*, a reduced exposure ($\sim 1/2$) direct contact print is provided in (a). Tracks (b-e) represent an identical exposure time. With *Mus*, identical hybridization patterns were observed with liver or total genomic DNA and the position and relative intensity of hybridizing components in Bam HI and Eco RI digests of total genomic DNA were identical to previous reports (11). Kb = kilobase, 3-day exposure.

corresponds to the abbreviated first hypervariable region but does not include the more extensive second and third hypervariable regions defined by S107V-F2. On prolonged exposure of the radioautographs, S107-F2 also detects an appreciable number of hybridizing components in the two phylogenetically distant species. The filter hybridization patterns with murine DNA and both subinserts were similar. S107V-A1, which is related to S107V-F1 but lacks the regions corresponding to the 5' untranslated and leader sequences, detected the majority of the stronger hybridizing components observed with the S107V-F1

Fig. 3 Cross-hybridization of *Sst* I-digested Mus (a), Tupaia (b) and Caiman (c) DNA with subinserts which were derived from S107V by digestion with *Fnu*4H1 (S107V-F1 [A] and S107-F2 [C]) or *Alu*I (S107V-A1 [B]) (Fig. 1). The specific activity of the nick-translated probe from which the subinserts were derived was 100 cpm/pg and individual hybridization mixtures contained 1.8 , 1.2 and 2.1×10^6 cpm for A, B and C respectively. The same distribution of hybridizable components in [B] was obtained when the probe bound to [A] was dissociated and rehybridized with S107V-A1, although the radioautographic intensity was somewhat diminished. All conditions were the same as in Fig. 2 except $5 \mu\text{g}$ of Mus DNA and a 10-day exposure were employed.

subinsert (Fig. 3), the subinsert corresponding to the untranslated and leader sequences hybridized only weakly with the Tupaia and Caiman sequences. A partial homology between the 5' probe sequence encoding V_H residues -3-49 and Tupaia and Caiman DNA sequences is suggested.

The hybridization patterns of $\mu 107V$, corresponding to the V_H region of $p\mu 107$, with *Sst* I-digested Mus, Tupaia, Gallus and Caiman genomic DNAs, are shown in Fig. 4. There were no components in Tupaia and Caiman DNA detected by $\mu 107V$ which were not detected with S107V and considering the relative specific activities of the probes (Figs. 2, 4), $\mu 107V$ apparently is superior for detecting cross-hybridizing components in these species. As with S107V, an apparent greater number of hybridizing components was visualized in Caiman

Fig. 4 Cross-hybridization of various vertebrate DNAs with μ 107V (13). Restriction mapping confirmed the identity of the probe which corresponds to UT, L, V, D, J (Fig. 1) and C μ 1 (2) genetic regions. The ~770nt probe was nick-translated to specific activity of 47 cpm/pg; all other conditions are identical to Fig. 2. Species order (a - e) and direct contact print exposure of Mus (a + b) as in Fig. 2.

vs Tupaia digests. The greatest difference was observed with Gallus and it is unlikely that the relative DNA content/cell in Mus, Tupaia, Gallus and Caiman (estimated for a related Crocodilian): 6.0, 6.6, 2.6, 5.0 μ g/ 10^6 cells, respectively, would explain the hybridization signals for the presumed single copy V_H genes. Furthermore, hybridization with the C μ 1 component of the probe would be unlikely since in mammals (22) and lower vertebrates (23), this region has been found to exhibit net reduced interspecies homology with few extended homologous segments and would not be encountered in as many individual hybridizing forms. Similarly, it is unlikely that D_H and J_H sequences are being detected. The extended 5' untranslated portion of μ 107V (2) may hybridize selectively with avian sequences, however, the similarity between S107V and μ 107V hybridization patterns in other species cannot be explained in this way. It also is possible that the selective recognition of avian sequences relates to differences in nucleic acid sequences in framework regions of the two

probes. When the nucleic acid sequence of $\mu 107V$ corresponding to the amino terminal 38 residues of the V region in the mature polypeptide (3) was compared to the corresponding sequence of S107V (1), 23% of the nucleotides were found to be different; excluding the first hypervariable region, the probes are 16% different at the nucleic acid level.

Taken together, the results suggest that the murine probes detect multiple homologous V_H sequences in several lower vertebrate species which were chosen on the basis of relative phylogenetic position (24). Tupaia, an insectivore, may represent the living mammal least changed from the ancestral placental mammals. While removed from the base of the reptilian tree, Caiman has sustained only moderate phenotypic change over an extended evolutionary time and shares a common Archosaur ancestor and morphologic features with Gallus, an avian. Ancestral forms of these latter two classes were present ~180-225 million years ago during the Triassic period. Efforts to extend these analyses to more phylogenetically distant species including Necturus, an amphibian, and Neoceratodus, a modern representative of a phylogenetically ancient lungfish form, utilizing the S107V and $\mu 107V$ probes, have not proven successful. Only weak hybridization confined to a single region and exhibiting less intensity than the ~4 kb component of Gallus illustrated in Figure 4 was observed in both cases (not illustrated). The high DNA content/cell (14) in these species may be a critical factor in detecting single copy gene products by the blotting procedure.

The diminished intensity of the bands (Figs. 2-4) relative to parallel analysis with homologous DNA and the need to utilize low stringency hybridization washing conditions suggest that appreciable sequence divergence at the nucleic acid level has occurred. In the absence of nucleic acid sequence data for the lower vertebrate hybridizable components, it is not possible to estimate the degree of homology; however, murine V_H probes have been shown to be capable of strong cross-hybridization with human V_H sequences, differing by 13% at the nucleic acid level when conditions of higher washing stringency are employed (7). Interpretation of the sequence regions being detected rests on the results of the experiments employing subinserts of S107V. Both S107V-F1 and S107V-A1 contain a portion of the V_H which exhibits interspecies homology at the amino acid level (23,25,26) and detects multiple hybridizing components in Caiman and Tupaia. To date amino acid sequencing of lower vertebrate V_H regions corresponding to the 3' portion of the V_H probe employed in these studies has not been carried out. In interpreting the apparent multiplicity in genetic regions which cross-hybridize with the murine V_H

probes, it should be noted that the use of nick-translated, relatively short probes facilitates identification of homologous individual gene segments which may not contribute to the pool of functional polypeptides (11,12).

The number of individual variable region genes is important in weighing the contributions of germline and somatic mechanisms of antibody diversity (3,5,11,12). With human (27) and mouse (5,28-30) κ chains, considerable variation exists in the relative extent of the germline commitment, while the number of V_H genes varies less extensively between these species (3,7). The results of these studies suggest that a complex V_H gene family may have an early evolutionary origin, although the absolute estimation of its extent would require analysis with homologous probes. Nevertheless, these studies demonstrate that a sufficient degree of homology may exist between V_H genes of advanced mammals and more phylogenetically distant forms to identify cross-hybridizing components in recombinant DNA libraries, thus making possible the identification of those DNA sequence features which reflect the origin, maintenance and stabilization of this multigenic family.

ACKNOWLEDGEMENTS

We thank Drs. L. Hood and T. Rabbitts for providing us with pS107V-1 and p μ 107, and Drs. M. Long and N. Williams for determining the DNA content/cell for Tupaia.

Supported by NIH CA-08748 and NSF-PCM-8001919.

REFERENCES

1. Early, P., Huang, H., Davis, M., Calame, K., and Hood, L. (1980) *Cell* 19, 981-992
2. Matthyssens, G., and Rabbitts, T.H. (1980) *Nucleic Acids Res.* 8, 703-713
3. Rabbitts, T.H., Matthyssens, G., and Hamlyn, P.H. (1980) *Nature* 284, 238-243
4. Brack, C., Hiram, M., Lenhard-Schuller, R., and Tonegawa, S. (1978) *Cell* 15, 1-14
5. Seidman, J.G., Leder, A., Nau, M., Norman, B., and Leder, P. (1978) *Science* 202, 11-17
6. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W., and Tonegawa, S. (1980) *Nature* 286, 676-683
7. Matthyssens, G., and Rabbitts, T.H. (1980) *Proc. Natl. Acad. Sci. USA* 77, 6561-6565
8. Kemp, D.J., Cory, S., and Adams, J.M. (1979) *Proc. Natl. Acad. Sci. USA* 76, 4627-4631
9. Siebenlist, U., Ravetch, J.V., Korsmeyer, S., Waldmann, T., and Leder, P. (1981) *Nature* 294, 631-635
10. Gough, N.M., and Bernard, O. (1981) *Proc. Natl. Acad. Sci. USA* 78, 509-513

11. Bothwell, A.L.M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K., and Baltimore, D. (1981) *Cell* 24, 625-637
12. Crews, S., Griffin, J., Huang, H., Calame, K., and Hood, L. (1981) *Cell* 25, 59-66
13. Blin, N., and Stafford, D.W. (1976) *Nucleic Acids Res.* 3, 2303-2308
14. Fasman, G.D. (1976) *Handbook of Biochemistry and Molecular Biology*, 3rd edn., Nucleic Acids - Vol II, CRC Press
15. Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517
16. Denhardt, D.T. (1966) *Biochem. Biophys. Res. Commun.* 23, 641-646
17. Jahn, C.L., Hutchison, C.A., Phillips, S.J., Weaver, S., Haigwood, N.L., Voliva, C.F., and Edgell, M.H. (1980) *Cell* 21, 159-168
18. Lederberg, E.M., and Cohen, S.N. (1974) *J. Bacteriol.* 119, 1072-1074
19. Maniatis, T., Jeffrey, A., and Kleid, D.G. (1975) *Proc. Natl. Acad. Sci. USA* 72, 1184-1188
20. Kataoka, T., Nikaido, T., Miyata, T., Moriwaki, K., and Honjo, T. (1982) *J. Biol. Chem.* 257, 277-285
21. Ben-Neriah, Y., Cohen, J.B., Rechavi, G., Zakut, R., and Givol, D. (1981) *Eur. J. Immunol.* 11, 1017-1022
22. Kehry, M., Sibley, C., Fuhrman, J., Schilling, J., and Hood, L.E. (1979) *Proc. Natl. Acad. Sci. USA* 76, 2932-2936
23. Litman, G.W., Erickson, B.W., Lederman, L. and Mäkelä, O. (1982) *Mol. Cell. Biochem.*, in press
24. Romer, A.S., and Parson, T.S. (1977) *The Vertebrate Body*, 5th edn., Holt, Rinehart and Winston, Inc., New York
25. Capra, J.D., and Kehoe, J.M. (1975) *Adv. Immunol.* 20, 1-40
26. Litman, G.W., and Marchalonis, J.J. (1982) in *Immune Regulation: Evolution and Biological Significance*, Ruben, L.N. and Gershwin, M.E. Eds., Marcel-Dekker, New York, in press
27. Bentley, D.L., and Rabbitts, T.H. (1981) *Cell* 24, 613-623
28. Cory, S., Tyler, B.M., and Adams, J.M. (1981) *J. Mol. Appl. Genet.* 1, 103-116
29. Selsing, E., and Storb, U. (1981) *Nucleic Acids Res.* 9, 5725-5735
30. Zeelon, E.P., Bothwell, A.L.M., Kantor, F., and Schechter, I. (1981) *Nucleic Acids Res.* 9, 3809-3820