# The genomic organization of immunoglobulin $V_H$ genes in *Xenopus laevis* shows evidence for interspersion of families

Robert N.Haire, Yuko Ohta<sup>+</sup>, Ronda T.Litman, Chris T.Amemiya<sup>§</sup> and Gary W.Litman<sup>\*</sup> University of South Florida, Department of Pediatrics, All Children's Hospital, 801 Sixth Street South, St Petersburg, FL 33701, USA

Received January 22, 1991; Revised and Accepted April 10, 1991

EMBL accession nos X56858-X56865 (incl.)

# ABSTRACT

The complete genomic sequences of Xenopus laevis immunoglobulin heavy chain variable region  $(V_H)$ genes comprising families IV – XI are reported. Using V<sub>H</sub> family-specific probes, linkage relationships for Xenopus  $V_H I - V_H XI$  have been determined. With the possible exceptions of  $V_HIII$  and  $V_HVII$ , Xenopus  $V_H$ genes appear to be interspersed. When from two to five V<sub>H</sub> segments are identified in individual clones, the elements are found to be in the same relative transcriptional orientation. Although the relationships of promoter sequences, including the regulatory octamer, resemble those seen in other vertebrate V<sub>H</sub> genes, several Xenopus V<sub>H</sub> families are associated with additional 5' octamer sequences and octamer-like motifs. The similarities between the genomic organization of V<sub>H</sub> genes in Xenopus and higher vertebrates, indicate an early phylogenetic emergence of the mammalian-type of gene organization and regulation.

# INTRODUCTION

A number of different approaches have been employed to study the developmental regulation of immunoglobulin gene expression in mammals (1-9). From these studies, it is apparent that the linear chromosomal organization of the rearranging segmental elements that encode the V<sub>H</sub> region is an important component in their utilization during development. Considerable insight into the mechanisms of B cell development and differentiation also has been gained from studies of lower vertebrates. In avians, clonal expansion of B lymphocytes in individual bursal follicles permits the unequivocal assignment of somatic gene conversion as an integral mechanism in diversifying the B cell response (10-12). Recently, this mechanism of somatic diversification has been established conclusively in a mammalian species (13). The amphibian represents another potentially valuable experimental model for studying the ontogenetic diversification of the B lymphocyte response (14). In *Xenopus laevis* (African clawed frog), the development of a larvae into an adult can be arrested by inhibiting thyroxine production with exposure to sodium perchlorate. This developmental block is associated with restricted structural (charge heterogeneity) and immunochemical (idiotype) differences in larval vs. adult *Xenopus* antibodies (15-18).

Immunoglobulin genes in *Xenopus* have been isolated either by DNA cross-hybridization employing murine DNA probes (19) or by immunochemical methods (20), and their structure and organization have been described. Preliminary analyses have shown the overall organization of the immunoglobulin V<sub>H</sub> locus of Xenopus (21) to be more like mammalian V<sub>H</sub> loci (22,23) than like the immunoglobulin genes of avians (12) or the more phylogenetically distant elasmobranchs (24,25). Recently, by analyzing immunoglobulin cDNA sequences we demonstrated that the complexity of  $V_H$  genes in *Xenopus*, in terms of the number of families, variation in complementarity determining region (CDR) and diversity (D<sub>H</sub>) region segments, and multiplicity of joining (J<sub>H</sub>) regions, approaches the complexity of murine immunoglobulin genes (26). Relatively little is known of the organization of the various families of V<sub>H</sub> genes in Xenopus. The three major families, I, II, and III, comprise about 70% of the total  $V_H$  elements isolated from a spleen cDNA library (26). Correspondingly, V<sub>H</sub>I, II and III constitute the majority of components identified by hybridization on a Southern blot of genomic DNA (26). The genomic sequences and organization of V<sub>H</sub>I, II, and III with respect to each other have been determined (27). The present study was undertaken to address the nature of the genomic structure and organization of  $V_{\rm H}$  genes in Xenopus in light of the eight additional  $V_{\rm H}$  gene families and to identify features of DNA sequence and organization that are potentially significant in terms of the regulation of the individual gene members.

<sup>\*</sup> To whom correspondence should be addressed

Present addresses: <sup>+</sup>Nagoya City University, Department of Pharmacology, Nagoya, Japan and <sup>§</sup>Lawrence Livermore Natl Lab, Div of Biomedical Sci L-452, Livermore, CA 94551, USA

## MATERIALS AND METHODS

# Probes

Probes for  $V_HI$ ,  $V_HII$  and  $V_HIII$  families were prepared by PCR using oligonucleotides complementing FR1 (nucleotides [nt] 1-18) and the family-specific regions of FR3 (excluding the highly conserved 3 terminus). Templates for the PCR probes were prototype or representative  $V_H$  family member cDNA clones. All other probes were described previously (26).

#### **Genomic library construction**

Genomic libraries were constructed in Lambda DASH<sup>TM</sup> (Stratagene) using standard procedures described previously by this laboratory (25). Two libraries were employed in these studies. The first was constructed using Sau3A digested DNA directly and yielded  $5 \times 10^5$  recombinants, with an average insert size of  $\sim 15-17$  kb, equivalent to  $\sim 1$  haploid genome. The second library was constructed using DNA that first was subjected to digestion with Sau3A and then fractioned by sucrose density centrifugation. Fractions corresponding to DNA in the 15-20 kb range were ligated to the  $\lambda$  DNA arms. Approximately  $5 \times 10^5$  recombinant phage, corresponding to  $\sim 1$  genome, were screened directly at moderate stringency, as described (28).

#### Mapping, subcloning and DNA sequencing

All clones initially identified with a  $V_H$ -specific probe were digested with restriction enzymes, Southern blotted and subsequently hybridized with each of the ten other  $V_H$  family-specific probes to identify additional  $V_H$  genes. Right and left arm-specific probes and analysis of partial digests facilitated mapping of restriction fragments. Subcloning in M13 mp18/19 used convenient restriction enzyme sites flanking the  $V_H$  genes. M13 (-40), 18 mer extension primers (as required with large subcloned fragments) and  $V_H$  family-specific FR1 (coding) and FR3 (reverse complement) primers were employed. DNA sequences were determined in both directions by the dideoxy method (29) using <sup>35</sup>S-dATP and T7 DNA polymerase, Sequenase<sup>TM</sup> (United States Biochemical Corporation). Sequences were compared with our database containing ~ 100 *Xenopus* cDNA sequences using GENALIGN<sup>TM</sup> (Intelligenetics).

#### Transcriptional orientation

Transcriptional orientation was determined following mapping using either a strand denaturation oligonucleotide hybridization method (30) or a novel method based on PCR. Intact phage clone templates were primed with combination FR1 and FR3 primers of differing polarity that complement adjacent genes of different V<sub>H</sub> families. Lambda arm-specific primers were substituted for one of the FR primers in some cases in which  $V_H$  genes were located near the ends of inserts. PCR products were analyzed on agarose gels, Southern blotted and hybridized with the full length family-specific probes. Orientation is deduced from the expected hybridization of the PCR fragments with two probes as well as from the predicted lengths of fragments based on the restriction map analysis. This PCR method was suitable for  $V_{H}$ segments separated by up to 7 kb. In a few cases, orientation was confirmed by DNA sequencing of fragments with asymmetric ends. Depending on the number of V<sub>H</sub> families and their relative proximity, transcriptional polarity in certain clones cannot be determined using either this method or the previously described technique (30).

## RESULTS

Multiple lifts of a genomic DNA library were screened with nine probes specific for Xenopus laevis V<sub>H</sub>-families V<sub>H</sub>II and  $V_{\rm H}IV - XI$  (26). The initial objective in this study was to isolate and sequence one genomic  $V_H$  gene representing families IV through XI. Positive-hybridizing plaques were selected solely on the basis of hybridization with a single V<sub>H</sub> family-specific probe, without consideration of hybridization intensity, which is influenced by the number of segments from an individual  $V_{H}$ family present in a single clone. Partial restriction map analyses of each of these isolates with all eleven  $V_H$  family probes indicate that all clones contain components that hybridize with probes specific for at least one gene family in addition to the  $V_{\rm H}$  family represented by the selection probe, with a single exception (see below). The sole criterion applied in selecting the gene segments that were subjected to DNA sequencing was the relative convenience of subcloning the component restriction fragment(s). The complete genomic sequences from a minimum of 300 and typically from 500-700 nt 5 of the initiation codon through at least the 3 recombination signal sequence (RSS) of a single, representative member of  $V_H$  families IV through XI have been submitted to the EMBL data base. EMBL Accession Numbers: X56858 (V<sub>H</sub>IV), X56859 (V<sub>H</sub>V), X56860 (V<sub>H</sub>VI), X56861 (V<sub>H</sub>VII), X56862 (V<sub>H</sub>VIII), X56863 (V<sub>H</sub>IX), X56864  $(V_HX)$  and X56865  $(V_HXI)$ .

In order to determine possible linkage relationships between  $V_H$  family members and between members of different  $V_H$ families, ten clones (average size, 17 kilobases [kb]) were subjected to partial restriction mapping (Fig. 1). Thirty-three V<sub>H</sub> gene segments were identified in the  $\sim 170$  kb mapped, giving an average distance of 5 kb between known elements. No evidence for overlap between individual clones is apparent. Two clones (XL 7; XL 4.6) contain two V<sub>H</sub> genes, five clones (XL 1.9; XL 1.2.10; XL 1.2.5; XL 1.8; XL 1.11b) contain three  $V_H$  genes, one clone (XL 2.3) contains four  $V_H$  genes and two clones (XL 1.8.11; XL 1.11a) contain five V<sub>H</sub> genes. Only one clone, XL 7, contains components belonging to a single family,  $V_H VII$ , and the two  $V_H VII$  genes differ by only four nt in their coding regions and by six nt and the length of the AT repeat in the 5 non-coding region. The three  $V_HXI$  genes appear to be identical over their amino acid coding regions. Two of these, the  $V_HXI$  genes from clones 1.11a and 1.11b, also are probably identical over their entire sequenced 5 portions (300 nt) and may represent allelic forms of the same DNA segment. These differ from the third gene, in clone 1.8.11, by only two nt in the 5 non-coding region. Three clones, XL 1.2.10; XL 1.2.5; XL 1.8.11, contain three different  $V_H$  families.

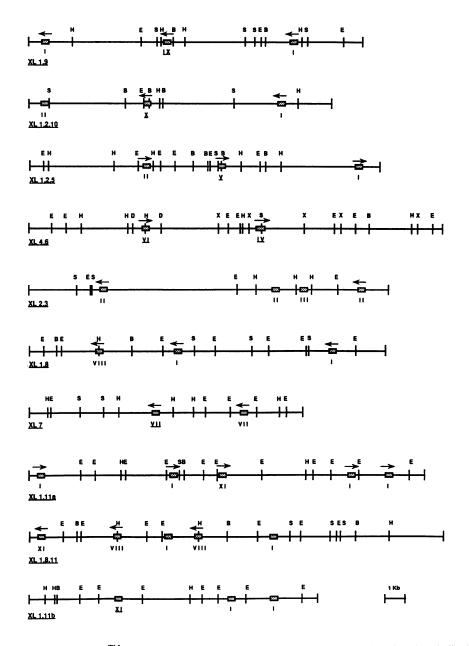
Based on the sequence data and polymerase chain reaction (PCR) priming of  $V_H$  families using family-specific oligonucleotide probes (see Materials and Methods), it is likely that each hybridizing segment represents an actual gene or closely related pseudogene. The 12 genes sequenced in the course of this study (see Fig. 1), however, do not appear to represent overt pseudogenes, based on analyses of coding regions, splice donor/acceptor and known functionally significant 5 and 3 sequence motifs. Fourteen of the 33 individual hybridizing gene segments found in these ten different  $\lambda$  clones hybridize with the V<sub>H</sub>I probe. Since V<sub>H</sub>I was not employed in the selection process, it is likely that this gene family is particularly abundant. The V<sub>H</sub>II family, considered one of the largest (26,27), is represented here by five elements, but is detected in only two

non-V<sub>H</sub>II selected clones. Similarly, V<sub>H</sub>III, another of the abundant families, appears only in the V<sub>H</sub>II selected clone XL 2.3. Recently, we have isolated 13 apparently different V<sub>H</sub>III clones and find only four of these contain members of another family, in all cases, V<sub>H</sub>II (work in progress).

While the number of genomic sequences that have been determined in this study is relatively small, comparison of these to the *Xenopus*  $V_H$  cDNA sequences identified by hybridization with  $J_{H^-}$  and  $C_H$ -specific probes (26) is informative. The conditions and selection criteria employed for specific hybridization require as little as ~60% nt identity, yet the identities between cDNA (probe) and genomic sequences range from 87% for  $V_HXI$  to 100% for one of the  $V_HVII$  genes (data

not shown). The complete identity between the predicted  $V_H$  coding segment of genomic  $V_HVII$  (XL 7 1.1 kb) and the previously reported sequence of a  $V_HVII$  cDNA (26) derived from an unrelated *Xenopus* specimen is remarkable and suggests that the corresponding  $V_H$  segment of this particular cDNA has not been altered somatically.

Although it is recognized that with the exception of  $V_HVII$ , the cDNA sequences previously reported (26) do not necessarily derive from these genomic sequences, comparisons reveal that most substitutions observed occur in CDRs. Of 122 total nt differences in the framework (FR) and CDR segments (boundary assignments correspond to those in [26]), ~45% are in CDRs which constitute only ~25% of the total sequence. The relative



**Figure 1.** Partial restriction maps of Lambda DASH<sup>TM</sup> clones.  $V_H$  genes are in boxes and the family designation is indicated, underlined family designations refer to genes sequenced and provided to EMBL. Genes sequenced but not presented in EMBL are the second  $V_HVII$  in XL7, both  $V_HXI$  genes from XL1.11a and XL1.8.11 and two  $V_HVII$  genes from clones not mapped in detail. All other  $V_H$  family designations were determined by hybridization. EcoRI (E), HindIII (H), BamHI (B), StI (S), DraI (D), and XbaI (X) sites are indicated by vertical lines. DraI and XbaI sites are shown only for XL 4.6, where they were used for subcloning. Average clone size is 17 kb. Arrows above the  $V_H$  genes indicate the relative transcriptional orientation (5–3). Clone number (underlined at left) is shown under each map and 1 kb scale is at the lower right.

rates of amino acid replacement in FR vs. CDR, normalized for different segment lengths (arbitrarily expressed here as amino acid changes/ten nt length), indicate that most amino acid changes are located in CDRs. Comparisons of the predicted amino acid sequences of the cDNA vs. the genomic clone isolated for each of the eight  $V_H$  families, reveal a range of 1–1.9 amino acid changes/ten nt for FR1, FR2 and FR3 with FR2 being the highest. The corresponding changes in CDR1 and CDR2 average 4.4 and 5.8 amino acid changes/ten nt, respectively.

The relative transcriptional polarity of  $V_H$  elements was examined in nine of the ten clones shown in Fig. 1. Because of the proximity of some  $V_H$  elements to the  $\lambda$  arms, extended linkage distances between certain  $V_H$  elements and the presence of more than one member of the same  $V_H$  gene family in some clones, it was not possible to determine transcriptional polarity of all of the elements; however, for those  $V_H$  genes in each clone that could be analyzed, the relative orientation is the same.

The distributions of immunoglobulin regulatory octamers (ATGCAAAT) (31,32) and octamer-motifs (defined as any sequences varying from the octamer consensus by no more than one nt) in the 5 upstream regions of the various genomic prototypes are illustrated in Fig. 2. At least one octamer is found  $\sim$  75–130 nt 5 of the initiation codon in representatives of each of the  $V_H$  families IV-XI. Previously it has been shown that Xenopus V<sub>H</sub>II genes possess a single octamer  $\sim 20$  nt 3 to the TATA sequence (27) and a  $V_{H}II$  gene with three perfect octamers within 600 bp of the initiation codon has been isolated (data not shown). Dual octamer sequences have been described in  $V_H$ I-type genes (27) and are present in the  $V_HV$  gene illustrated in Fig. 2. Only the V<sub>H</sub>III-type genes appear to lack perfect octamer sequences (27). Additional octamer-motifs are located various distances 5 of the ATG initiation codon in Xenopus V<sub>H</sub>V, VI, VII, VIII, IX and×gene families; however, the functional significance of these structures is not understood. With  $V_HV$ , VI, IX and X, the *octamer*-motifs are located 5 of the ATGCAAAT; the octamer-motif present in the V<sub>H</sub>VII gene is located 3 to the typical octamer.

In these studies, the potential immunoglobulin TATA box is defined operationally as any T/A-only sequence stretch of at least five nt that: 1) initiates with TA..., 2) consists of at least two Ts and three As or two As and three Ts or is the sequence TAAAA and 3) is located downstream of an octamer or *octamer*-

motif. By these definitions, each of the sequences reported here possesses at least one TATA box. Leader sequences typically consist of 18 amino acids, with the exception of the 19 amino acids in the leader segment of V<sub>H</sub>VIII genes. The leader intervening sequences (IVS) found in different families of  $V_H$ genes exhibit little or no sequence identity. Overall nucleotide sequence relatedness within the extended 5 noncoding region of particular V<sub>H</sub> genes selected in this study is also very low, with the exception of the octamer, octamer-motif and TATA boxes. Some extended regions of sequence identity, however, are shared by pairs of gene families. A nonamer sequence, CAATGCC-TT, is located 5 of the octamer in the  $V_H IV$  and  $V_H V$  genes. An extended region of sequence identity, TAATATATG, N<sup>16</sup> or N<sup>19</sup>, TAAATACATT is located  $\sim 200$  nt upstream of the most 5 octamer-motif in the otherwise unrelated V<sub>H</sub>VII and V<sub>H</sub>IX genes. A few additional eight-nt sequences also are shared. V<sub>H</sub>X and XI share ATTCTCAT at 131 and 41 nt, respectively, upstream of the octamers. ATTTGCAT, the octamer inverse complement that functions in the regulation of immunoglobulin light chain gene expression, was not detected in the 5 untranslated region of these genes, although a perfect inverse octamer and an inverse octamer-motif are located 3 of the V<sub>H</sub>XI and V<sub>H</sub>VII genes, respectively. Potential regulatory roles, if any, of the other shared elements are unknown.

A heptamer-nonamer RSS is located 3 of the coding segment of each gene. Representative members of eight of the eleven  $V_H$ families possess the prototypic sequence CACAGTG (27). A single nt difference at the fifth position of the heptamer is found in  $V_H$ VIII, IX and XI. The nonamer sequences exhibit considerably more variation. The 23 nt  $V_H$  RSS spacers exhibit ~40-50% identity. The identity in the spacer sequences between gene families suggests functional significance beyond that of simple physical separation (33). Conserved identities in RSS spacers have been noted in species representing more distant levels of phylogenetic development (28), as well as in more phylogenetically recent species (34).

## DISCUSSION

Unique aspects of the ontogeny of the antibody response in *Xenopus laevis* make this a potentially valuable model for understanding the developmental regulation of immunoglobulin

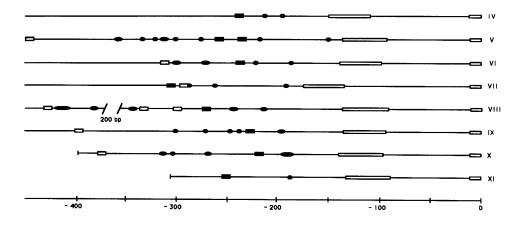


Figure 2. Position of 5 elements relative to the  $V_H$  gene leader sequences. The split leader sequences are at the right occupying about 150 nt (open rectangles) and ending at the 5 end of the mature immunoglobulin (position 0). TATA-like sequences are solid ovals, perfect octamers are filled boxes and imperfect octamers are open boxes. Certain sequences for which we have additional 5 DNA sequence information are shown without a vertical line at the left.

gene expression. Recent efforts have resulted in the cloning and sequencing of a Xenopus V<sub>H</sub>I gene (19), an IgM-type gene (20) as well as genes encoding two other immunoglobulin heavy chain classes IgX (35) and IgY (36). An additional family of V<sub>H</sub> genes has been identified and linkage relationships of V<sub>H</sub>I, II and III genes have been examined (27). One report concluded that the V<sub>H</sub> repertoire in *Xenopus* is restricted in terms of both the number of  $V_H$  families and the extent to which these are modified somatically (27). Another report indicates that this species contains at least five  $V_H$  families (37). More recently, using an iterative screening procedure, we isolated and classified 200 cDNA clones in 11 distinct  $V_H$  families that vary as much, if not more, from one another than comparable  $V_H$  families in higher vertebrates (26). Furthermore, high degrees of variation in CDR and  $D_H$  sequence segments as well as heterogeneity in  $J_{H}$  segments are apparent. In order to better understand the origins of gene diversity in this species, a study of the genomic organization of immunoglobulin V<sub>H</sub> genes was undertaken.

Previously, we reported that differences in complex patterns observed on genomic Southern blotting with V<sub>H</sub> family-specific probes are consistent with the presence of distinct, individual  $V_H$ families (26). Comparisons of these previously reported cDNA sequences to the genomic sequence data presented here indicate a maximum DNA sequence difference of only 13%, confirming that the putative  $V_H$  families do not arise by intergenic recombination. Variation within a V<sub>H</sub> family can be as high as 12 nt in V<sub>H</sub>IV FR1, representing a 10% difference, and seven nt in  $V_HXI$  FR2, representing a 16% difference. These appreciable differences suggest that subsets of  $V_H$  families may exist because high levels of somatic modification of FR segments during B cell development are not likely. The DNA (coding) sequence of one of the genomic V<sub>H</sub>VII elements is identical to a previously reported V<sub>H</sub>VII cDNA. Both genomic V<sub>H</sub>VII genes from clone XL 7 code for Tyr instead of the nearly universal Trp located at the first residue of FR2 in the V<sub>H</sub> genes of all other species. Only two mouse and one human gene out of almost 500 summarized differ at this position (38).

The lack of appreciable identity in 5 regions, with the exceptions noted, is further indicative of the presence of distinct  $V_{\rm H}$  gene families in *Xenopus*. The unique configuration of octamer elements in some of the  $V_H$  families may be of functional significance and could be of particular developmental consequence in this species. Also of note are the octamer-motif sequences that are associated with certain V<sub>H</sub> families. Although evidence has been presented that a perfect octamer sequence is required for full promoter function, certain substitutions can diminish and perhaps alter the specificity of this promoter element (39-41). Such sequence features could be significant in terms of the developmental regulation of  $V_H$  gene expression in this species. Recently, octamer-motifs have been implicated in the replicative transposition of another Xenopus gene (42). The presence of additional 5 extended sequence identities between some of the gene families may be important in terms of understanding coordinate utilization and the phylogenetic diversification of distinct gene families.

In the mouse, immunoglobulin  $V_H$  genes belonging to a family tend to be arranged in closely linked clusters; however, a few families are dispersed within the locus (43,44). Ordered expression of murine immunoglobulin genes correlates with chromosomal position (3,45,46), although evidence has been presented that the expression may be regulated at a different locus (3,47,48). In contrast, individual human immunoglobulin  $V_H$ 

genes belonging to different families are interspersed (22,23). Interspersion of V<sub>H</sub> families also has been observed in phylogenetically primitive teleost fish (49) (Amemiya and Litman, unpublished observations). Assuming that the different observations regarding linkage are not due to differences in methodology, i.e., the human and *Xenopus* data are based on analyses of cloned genomic fragments, whereas the mouse data are based largely on deletion mapping during DNA rearrangements, it appears that major differences in V<sub>H</sub> gene organization have arisen in vertebrate evolution.

Originally, a genomic clone with four hybridizing components, presumably representing V<sub>H</sub>1-type genes, was detected in a Xenopus genomic DNA library by cross-hybridization (19). In the absence of other  $V_H$  family-specific probes, it was not possible to ascertain whether this clone contained members of additional  $V_H$  families. More recently, the chromosomal order  $V_HI-V_HII-V_HIII$  was proposed (27). The results reported here for 11 Xenopus V<sub>H</sub> families, including V<sub>H</sub>I-V<sub>H</sub>II-V<sub>H</sub>III, indicate that V<sub>H</sub>I genes are interspersed with the members of six other  $V_H$  families. Given the very large numbers of  $V_H$  genes described thus far and the mapping data reported here, it is possible to refine the known organization of the gene locus. Most likely,  $V_{H}I$  and to a lesser extent  $V_{H}II$  elements are highly interspersed with members of the other  $\boldsymbol{V}_{\boldsymbol{H}}$  families. Members of one of the most abundant V<sub>H</sub> gene families, V<sub>H</sub>III, are apparently not found in close linkage with members of other gene families except for limited interspersion with V<sub>H</sub>II members. It is more difficult to conclude whether the V<sub>H</sub>VII genes are interspersed, since this gene family is less complex and only a single ( $V_H VII$ ) clone has been identified and characterized.

Although transcriptional polarity is an integral component in the regulation of eukaryotic gene expression, relatively little is known regarding the relative orientation of immunoglobulin heavy chain genes. A recently described, denaturation-based technique (30) and a method employing PCR (see Materials and Methods) provide unequivocal information about relative orientation of genes within the Xenopus genomic V<sub>H</sub> clones. In the physical maps of nine of the individual clones presented in Fig. 1, the transcriptional orientation of immunoglobulin genes is the same. While it is not possible to conclude that all  $V_{H}$ genes in Xenopus are oriented in the same manner, individual  $V_{\rm H}$  segments representing all eleven families share the same transcriptional polarity relative to closely linked members of different families. It remains possible that other genes, perhaps those exhibiting different patterns of developmental expression, may possess different transcriptional orientations.

The  $V_H$  gene system of *Xenopus* is remarkably similar to that of mammals in terms of genomic complexity. The studies reported here provide evidence that the separate V<sub>H</sub> elements constituting the  $V_H$  locus are interspersed, similar to findings reported for human and some murine V<sub>H</sub> families. Interspersion of V<sub>H</sub> genes probably represents the most effective form of organization in terms of non-preferential utilization of genes. The principal exception involves the V<sub>H</sub>III-type (and V<sub>H</sub>VII-type, see above) genes, that were not detected in association with  $V_HIV - XI$ . In Xenopus,  $V_HIII$  appears to be utilized preferentially in early development (R. Haire unpublished observations); non-interspersion may represent an effective means for preferred utilization of this gene family through cis- and/or trans-mechanisms. It should be noted that the Xenopus V<sub>H</sub>III family is not closely related in nucleotide sequence to mammalian V<sub>H</sub>III genes which are preferentially expressed in fetal mice (45) and humans (5). The analyses of both gene organization and 5 flanking sequence structure reported here are essential in the next phase of studies which seek to identify the mechanisms of selective utilization of  $V_{\rm H}$  genes in early development.

#### ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health grant GM-38656. Dr. R.N.Haire is the recipient of a fellowship from the American Cancer Society. Dr. G.W.Litman is the Hines Professor at the University of South Florida College of Medicine/All Children's Hospital. We thank Mrs. Barbara Pryor for editorial assistance.

#### REFERENCES

- Alt, F.W., Rosenberg, N., Enea, V., Siden, E. and Baltimore, D. (1982) Mol. Cell. Biol. 2, 386-400.
- Blackwell, T.K. and Alt, F.W. (1989) Ann. Rev. Genet. 23, 605-636.
   Yancopoulos, G.D., Desiderio, S.V., Paskind, M., Kearney, J.F., Baltimore,
- D. and Alt, F.W. (1984) Nature 311, 727-733.
  4. Perlmutter, R.M., Kearney, J.F., Chang, S.P. and Hood, L.E. (1985) Science
- 227, 1597 1601.
   Schroeder, H.W. and Wang, J.Y. (1990) Proc. Natl. Acad. Sci. USA 87,
- 5. Schroeder, H. W. and Wang, J. I. (1990) Proc. Natl. Acad. Sci. USA 8. 6146-6150.
- 6. Jeong, H.D. and Teale, J.M. (1988) J. Exp. Med. 168, 589-603.
- 7. Schulze, D.H. and Kelsoe, G. (1987) J. Exp. Med. 166, 163-172.
- 8. Wu, G.E. and Paige, C.J. (1986) EMBO J. 5, 3475-3481.
- 9. Gu, H., Forster, I. and Rajewsky, K. (1990) EMBO J. 9, 2133-2140.
- Reynaud, C.-A., Anquez, V., Dahan, A. and Weill, J.-C. (1985) Cell 40, 283-291.
- Reynaud, C.-A., Anquez, V., Grimal, H. and Weill, J.-C. (1987) Cell 48, 379-388.
- Reynaud, C.-A., Dahan, A., Anquez, V. and Weill, J.-C. (1989) Cell 59, 171-183.
- 13. Becker, R.S. and Knight, K.L. (1990) Cell 63, 987-997.
- Du Pasquier, L., Schwager, J. and Flajnik, M.F. (1989) Ann. Rev. Immunol. 7, 251-275.
- 15. Hsu, E. and Du Pasquier, L. (1984) Differentiation 28, 109-115.
- 16. Du Pasquier, L. and Blomberg, B. (1982) Immunogenetics 15, 251-260.
- 17. Du Pasquier, L. and Wabl, M.R. (1978) Eur. J. Immunol. 8, 428-433.
- 18. Hsu, E. and Du Pasquier, L. (1984) Differentiation 28, 116-122.
- 19. Yamawaki-Kataoka, Y. and Honjo, T. (1987) Nucleic Acids Res. 15, 5888.
- 20. Schwager, J., Mikoryak, C.A. and Steiner, L.A. (1988) Proc. Natl. Acad.
- Sci. USA 85, 2245-2249. 21. Schwager, J., Grossberger, D. and Du Pasquier, L. (1988) EMBO J. 7,
- 2409-2415.
  22. Kodaira, M., Kinashi, T., Umemura, I., Matsuda, F., Noma, T., Ono, Y. and Honjo, T. (1986) J. Mol. Biol. 190, 529-541.
- Berman, J.E., Mellis, S.J., Pollock, R., Smith, C.L., Suh, H., Heinke, B., Kowal, C., Surti, U., Chess, L., Cantor, C.R. and Alt, F.W. (1988) EMBO J. 7, 727-738.
- 24. Hinds, K.R. and Litman, G.W. (1986) Nature 320, 546-549.
- Kokubu, F., Litman, R., Shamblott, M.J., Hinds, K. and Litman, G.W. (1988) EMBO J. 7, 3413-3422.
- Haire, R.N., Amemiya, C.T., Suzuki, D. and Litman, G.W. (1990) J. Exp. Med. 171, 1721–1737.
- Schwager, J., Burckert, M., Courtet, M. and Du Pasquier, L. (1989) EMBO J. 8, 2989-3001.
- Litman, G.W., Berger, L., Murphy, K., Litman, R., Hinds, K.R. and Erickson, B.W. (1985) Proc. Natl. Acad. Sci. USA 82, 2082-2086.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- Roditi, I., Konig, E. and Williams, R.O. (1989) Nucleic Acids Res. 17, 10506.
- Parslow, T.G., Blair, D.L., Murphy, W.J. and Granner, D.K. (1984) Proc. Natl. Acad. Sci. USA 81, 2650-2654.
- 32. Falkner, F.G. and Zachau, H.G. (1984) Nature 310, 71-74.
- 33. Tonegawa, S. (1983) Nature 302, 575-581.
- Schroeder, H.W., Hillson, J.L. and Perlmutter, R.M. (1990) International Immunol. 20, 41-50.

- Haire, R., Shamblott, M.J., Amemiya, C.T. and Litman, G.W. (1989) Nucleic Acids Res. 17, 1776.
- Amemiya, C.T., Haire, R.N. and Litman, G.W. (1989) Nucleic Acids Res. 17, 5388.
- Hsu, E., Schwager, J. and Alt, F.W. (1989) Proc. Natl. Acad. Sci. USA 86, 8010-8014.
- Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H.M. and Gottesman, K. (1987). Sequences of Proteins of Immunological Interest, U.S. Dept. Health and Human Services, Washington, D.C.
- Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P.A. and Baltimore, D. (1986) Nature 323, 640-643.
- 40. Wirth, T., Staudt, L. and Baltimore, D. (1987) Nature 329, 174-178.
- 41. Atchison, M.L., Delmas, V. and Perry, R.P. (1990) EMBO J. 9, 3109-3117.
- Morgan, G.T. and Middleton, K.M. (1990) Nucleic Acids Res. 18, 5781-5786.
- Brodeur, P.H., Osman, G.F., Mackle, J.J. and Lalor, T.M. (1988) J. Exp. Med. 168, 2261-2278.
- 44. Rathbun, G.A., Capra, J.D. and Tucker, P.W. (1987) EMBO J. 6, 2931-2937.
- Malynn, B.A., Yancopoulos, G.D., Barth, J.E., Bona, C.A. and Alt, F.W. (1990) J. Exp. Med. 171, 843-859.
- Yancopoulos, G.D., Malynn, B.A. and Alt, F.W. (1988) J. Exp. Med. 168, 417-435.
- 47. Wu, G.E. and Paige, C.J. (1988) J. Exp. Med. 167, 1499-1504.
- Jeong, H.D., Komisar, J.L., Kraig, E. and Teale, J.M. (1988) J. Immunol. 140, 2436-2441.
- Amemiya, C.T. and Litman, G.W. (1990) Proc. Natl. Acad. Sci. USA 87, 811-815.