

Conserved structure of amphibian T-cell antigen receptor β chain

(T-cell receptor/urodele amphibian/evolution)

JULIEN S. FELLAH, FABIENNE KERFOURN, FRANCOISE GUILLET, AND JACQUES CHARLEMAGNE

Université Pierre et Marie Curie and Centre National de la Recherche Scientifique, Unité de Recherche Associée 1135, Immunologie Comparée, 9 quai Saint Bernard, 75252 Paris Cédex 05, France

Communicated by Max D. Cooper, April 16, 1993

ABSTRACT All jawed vertebrates possess well-differentiated thymuses and elicit T-cell-like cell-mediated responses; however, no surface T-cell receptor (TCR) molecules or TCR genes have been identified in ectothermic vertebrate species. Here we describe cDNA clones from an amphibian species, *Ambystoma mexicanum* (the Mexican axolotl), that have sequences highly homologous to the avian and mammalian TCR β chains. The cloned amphibian β chain variable region (V β) shares most of the structural characteristics with the more evolved vertebrate V β and presents \approx 56% amino acid identities with the murine V β 14 and human V β 18 families. The two different cloned axolotl β chain joining regions (J β) were found to have conserved all the invariant mammalian J β residues, and in addition, the presence of a conserved glycine at the V β -J β junction suggests the existence of diversity elements. The extracellular domains of the two axolotl β chain constant region isotypes C β 1 and C β 2 show an impressively high degree of identity, thus suggesting that a very efficient mechanism of gene correction has been in operation to preserve this structure at least from the early tetrapod evolution. The transmembrane axolotl C β domains have been less well conserved when compared to the mammalian C β but they do maintain the lysine residue that is thought to be involved in the charged interaction between the TCR $\alpha\beta$ heterodimer and the CD3 complex.

A fundamental question in the evolution of the vertebrate immune system is the origin of the T-cell- and B-cell-specific antigen receptors. Mammalian and avian T lymphocytes recognize foreign molecules by using T-cell receptors (TCRs), disulfide-linked TCR $\alpha\beta$ and TCR $\gamma\delta$ heterodimers (for review, see refs. 1–3). Each TCR α , β , δ , and γ polypeptide chain is composed of two immunoglobulin-like constant (C) and variable (V) domains and the diversity of V domains is achieved, as for immunoglobulin molecules, by the use of different combinations of V, diversity (D), and joining (J) segments and by junctional diversity during the rearrangement of these segments (for review, see ref. 4). The TCR $\alpha\beta$ heterodimer (TCR2) seems largely responsible for the recognition of processed peptides by cytotoxic and helper T cells in the context of the major histocompatibility (MHC) class I and class II molecules. The TCR $\gamma\delta$ heterodimer (TCR1) is mostly expressed by intraepithelial T cells and would recognize antigens in either a non-MHC- or MHC-restricted manner (4).

All the actually studied representative species of nonamniotic jawed vertebrates possess well-differentiated thymuses and elicit efficient immune responses that would be in part, by analogy with the mammalian models, mediated or regulated by thymus-derived cells. However, T-cell-dependent responses have been experimentally demonstrated in a very limited number of species (5), and until now,

surface TCR-like molecules or TCR genes have not been identified in ectothermic vertebrate species.

By starting from the consensus opinion that immunoglobulin and TCR genes may be derived from a common ancestral gene (6), our interest focused on an experimental strategy that could help characterize TCR genes in ectothermic vertebrates. Our thought started from the recently published sequence analysis of a primitive Chondrichthyan fish (*Heterodontus francisci*) light chain (7). The *Heterodontus* λ chain constant (C λ)-like chains present important regions of nucleotide sequence identity with the avian and mammalian TCR β chains and an extensive sequence comparison of all known vertebrate λ light chains and TCR β chains confirmed the presence of long stretch of conserved nucleotides in the N-terminal one-third of the C λ and C β domains. A 32-fold degenerated 29-mer oligonucleotide corresponding to an area of this conserved region was designed, which allowed the cloning and structural analysis of cDNAs encoding the complete TCR β chain of an urodele amphibian species, the Mexican axolotl.*

MATERIALS AND METHODS

Axolotls. Neotenic axolotls of the Ax6 strain were bred in our laboratory colony in 14–16°C tap water. Immunizations of adults were carried with trinitrophenol-coupled sheep erythrocytes as described (8).

Isolation and Characterization of Axolotl TCR β cDNA Clones. Total RNA (1 μ g) from axolotl thymocytes was used for the first-strand cDNA synthesis in the presence of the 3'-end first-strand primer 5'-GACTCGAGTCGACATCGA-T₁₇-3', as described (9). The amplification step was carried out in 30 μ l of reaction mixture containing \approx 1% of the cDNA-RNA hybrid, the above 3'-end primer (2–4 μ g/ml), the TCR β primer [5'-GGGAATTCAAGGC(T/C)AC(C/A)CT(G/C)GTGTG(C/T)(T/C)TG-3'] (2–4 μ g/ml), and 1 unit of *Taq* polymerase (Cetus) in 10 mM Tris-HCl, pH 8.3/50 mM KCl/1.5 mM MgCl₂/10 mM 2-mercaptoethanol/0.01% gelatin. PCR was carried out in a Hybaid machine at 94°C for 0.6 min, 50°C for 0.6 min, and 72°C for 2 min for 30 cycles followed by 9 min at 72°C (10). Amplified DNA (600 bp) was purified (GeneClean) and cloned in the *Sma* I site of the Bluescript vector (Stratagene). One of the clones (T600.8) was sequenced in both strands by the dideoxynucleotide method (11) using universal primers and synthetic oligonucleotides and corresponds to the major part of C β from aa 150 (12) to aa 288, followed by a stop codon (TGA) and the complete 3' untranslated (UT) segment. An axolotl spleen cell λ ZAP II cDNA library (13) was screened by hybridization with a 390-bp random-primed ³²P-labeled probe obtained by PCR amplification of clone T600.8 using the above TCR β

Abbreviations: TCR, T-cell receptor; V, variable; D, diversity; C, constant; J, joining; V β , C β , etc., β chain V region, β chain C region, etc.; UT, untranslated; EC, extracellular; TM, transmembrane.

*The sequences reported in this paper have been deposited in the GenBank data base (accession nos. X70168 and L08498).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

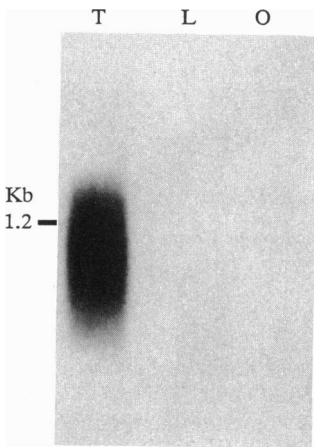


FIG. 1. Northern blot analysis of TCR β gene expression. Total RNA from axolotl thymocytes (lane T), liver (lane L), and ovaries (lane O) was hybridized to a ³²P-labeled probe encoding the C β 1 region.

primer and a reverse primer complementary to the 3' end of the C β 1 coding sequence (5'-TCACAGCTTTACTCTGCACAT-3'). Positive clones were purified and excised *in vivo* with the M13 VCS phage (Stratagene).

RNA Blot Analysis. Total thymus, liver, and ovary RNA was denatured, fractionated on agarose formaldehyde gels, transferred to nitrocellulose as described (13), and hybridized with the C β -specific ³²P-labeled random-primed probe (see above). After washing, the Northern blot was autoradiographed at -80°C for 1-3 days.

RESULTS AND DISCUSSION

Characterization of cDNA Clones. A survey of the vertebrate sequences encoding the C regions of the TCR β and immunoglobulin C λ chains indicates a strongly conserved stretch of nucleotides around the first N-terminal cysteine residue, which prompted us to design a moderately (32-fold) degenerated 29-mer oligonucleotide primer that allowed the amplification by RNA/PCR (9) of a 600-bp cDNA segment from axolotl thymocyte RNA. This cDNA (clone T600.8) encodes a putative polypeptide that has significant homologies with both the mammalian and avian C β chains (data not shown). This cDNA was used to screen an axolotl spleen cell cDNA library and several clones were selected. Clones 24 and 1⁺ include 5'-truncated V β segments followed by identical C β 1 segments and complete 109-nt 3'-UT sequences. Clone 22⁺(1) encodes a 5'-truncated leader segment followed by a complete V β domain, a C β domain that represents a second C β 2 isotype, and an almost complete 3'-UT 272-nt sequence.

Lymphoid Tissues Specifically Express C β mRNA. A Northern blot of total cellular RNA from axolotl thymocytes, liver, and ovary was hybridized with a ³²P-labeled axolotl C β probe. A strong hybridization signal at \approx 1.2 kb was obtained with thymocyte RNA but none was detected with ovary and liver RNA (Fig. 1).

Structure of the Axolotl V β . A search in the sequence data base (GenBank Release 73) using the FASTA program (14)

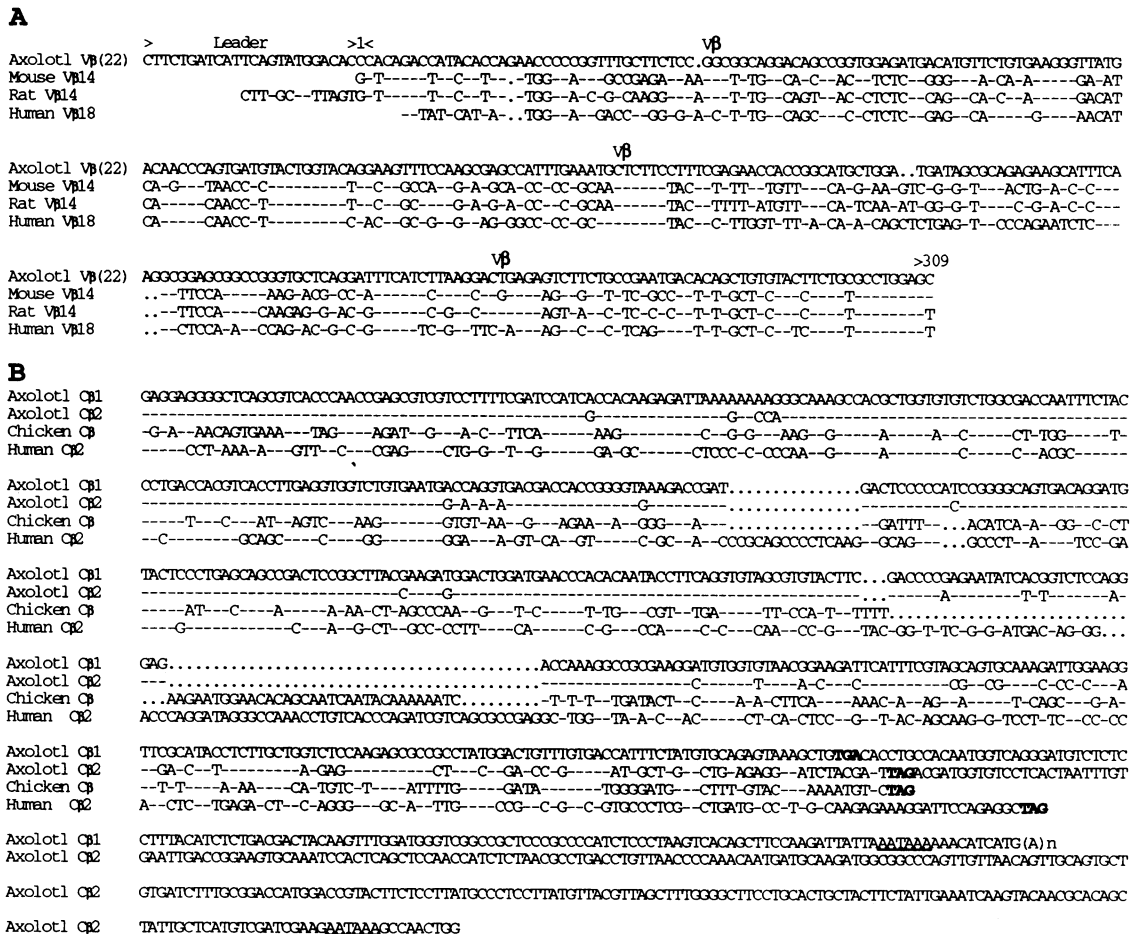


FIG. 2. (A) Nucleotide sequence of the axolotl cDNA clone 22⁺(1) V β aligned to rodent V β 14 and human V β 18 segments. The overall nucleotide identity between the axolotl V β and murine V β 14 regions is \approx 56%. (B) Nucleotide sequence homology of the axolotl C β 1 (clone 1⁺) and C β 2 [clone 22⁺(1)] regions and the equivalent chicken C β (3) and human C β 2 (18) regions. Stop codons are in boldface type, the 3'-UT sequences are indicated only for the axolotl genes, and the putative polyadenylation signal sequences are underlined. In A and B, residues identical to the axolotl V β and C β 1 sequences are denoted by hyphens and spaces, introduced to optimize homology between coding sequences, are indicated by dots.

```

>          Vβ          < N Gly Dβ N <          Jβ          >
          Cys          Gly          PheGly Gly Leu
24      GTATGAGTTTGTAAATGCGGTGGGACAGGGTCTCTACGGACGCTCTGCTTGGCCAGGGGACAAAACCTACCGTGTCTA/Cβ1
1*      -----
22*(1)  --G-ACT-C--CGCCTGGAGCATCTTC--GAC.....CAGTA-----AG-A--A-C--G--C--T--CG-C/Cβ2
    
```

FIG. 3. Nucleotide sequence of the 5' end of cDNA clones 1⁺ and 22⁺(1) compared to the equivalent sequence of clone 24. Above the sequences are indicated the putative limits of the Vβ, Dβ, and Jβ regions and the conserved amino acid residues.

revealed significant similarity scores (150 < opt < 250) between the axolotl Vβ22⁺(1) segment and the aligned Vβ sequences belonging to the rodent Vβ14 family (15, 16) and the corresponding human Vβ18 family (17) (Fig. 2A). At the amino acid level, the axolotl Vβ matches the mammalian Vβ for most of the characteristic residues (see Fig. 4A) including the invariant amino acids Gln-6, Pro-8, Cys-23, Trp-34, Tyr-35, and Cys-92, according to the Kabat numbering system (12). In rodent Vβ14 and human Vβ18, positions 62A, 63, and 64 are lacking and a leucine residue is found in position 65, thus leading to some difficulty to classify these families according to the Kabat Vβ1/VβII subgroups (19). In axolotl, a putative salt bridge could form between His-62 and Asp-86 and a phenylalanine residue is found in position 65. The Jβ amino acid sequence of clone Vβ22⁺(1) conserved the four mammalian Jβ invariant residues (Phe-108/Gly-109/Xaa/Gly-111/Xaa/Xaa/Leu-114) (12). Clones 24 and 1⁺ use identical Jβ regions, different from the clone Vβ22⁺(1) Jβ, and the Vβ 3' end of clone 24 is clearly different from the corresponding segment in Vβ22⁺(1) (Fig. 3). The conserved glycine residue in the junctional area between the Vβ and Jβ regions and the different sizes of these areas provide evidence for Dβ-like segments and for junctional hypervariability in the axolotl Vβ complementarity-determining region 3.

Structure of the Axolotl Cβ1 and Cβ2 Regions. The axolotl Cβ1 sequence was compared with the NBRF sequence data base (Release 35) by using the FASTA alignment program (14).

Axolotl Cβ1 is most closely related to the chicken Cβ (opt = 317) followed by various mammalian Cβ1/Cβ2 sequences (217 < opt < 267). Significant scores were also found with mammalian Cλ and Cκ chains, with the chicken Cλ chain, and with Chondrichthyan Cλ-like sequences. Extensive nucleotide sequence homology can be seen between large areas of the axolotl Cβ1 and Cβ2 genes (Fig. 2B). Of the first 100 codons at the 5' end, 88 are completely conserved and 2 present synonymous substitutions. Such high conservation between genes that duplicated ≈130 million years B.P., before the Ambystomatidae speciation (20), must be of survival importance and cannot be accounted for merely by the selective environmental pressure. Therefore, it would seem that there exists a gene correction mechanism, such as gene conversion, similar to that suggested for mammalian immunoglobulin and Cβ genes (21–23). The highly divergent 3'-UT and transmembrane sequences of the Cβ1 and Cβ2 cDNAs rule out the possibility that the corresponding genes are highly diversified Cβ alleles. The axolotl Cβ sequences encode constant regions of 154 (Cβ1) or 156 (Cβ2) aa (Fig. 4B) that can be compared to the chicken Cβ (157 aa) but are significantly shorter than mammalian Cβ (179 aa in humans). Most of this size difference is located at the level of the connective peptide, which is 14 aa shorter in the axolotl. In the extracellular (EC) domain, 39 residues (≈40%) are conserved in the axolotl, avian, and mammalian Cβ chains, including most of the residues considered to be important for

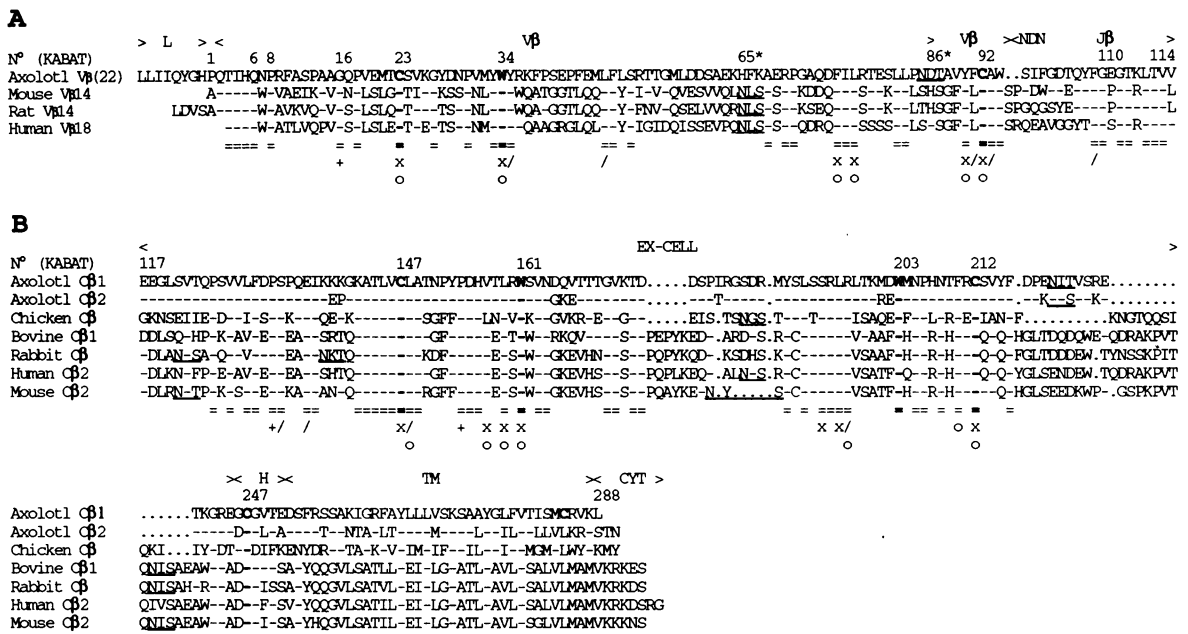


FIG. 4. (A) Amino acid sequence homology of axolotl Vβ [clone 22⁺(1)], rodent Vβ14 (15, 16), and human Vβ18 (17) domains. The putative limits of leader (truncated), Vβ, Dβ, and Jβ segments are indicated above the sequences. (B) Amino acid sequence homology of the Cβ1 and Cβ2 axolotl regions and chicken (3), bovine (24), rabbit (25), human (18), and mouse (26) Cβ regions. The sequence is divided into four sections, according to the corresponding known mammalian exons (18, 26) encoding the presumptive EC (EX-CELL), hinge (H), TM, and cytoplasmic (CYT) regions. In A and B, residues identical to the axolotl Vβ and Cβ1 sequence are denoted by hyphens and spaces, introduced to optimize homology between sequences, are indicated by dots. The conserved cysteine and tryptophan residues are in boldface type and sites of potential N-linked glycosylation are underlined. Important conserved amino acids are numbered according to Kabat *et al.* (12), and in the Vβ sequence, asterisks indicate the His-62 and Asp-86 residues that could form a putative salt bridge (19). Notations below the sequence signify characteristic amino acids: =, conserved in all sequences; x, β-strand characteristic residues at positions where the side chains point inside the sheet; o, hydrophobic residues thought to be buried in the domain core; +, two proline residues considered to be important in the architecture of loops; /, residues occupying key positions in domain–domain contacts (27).

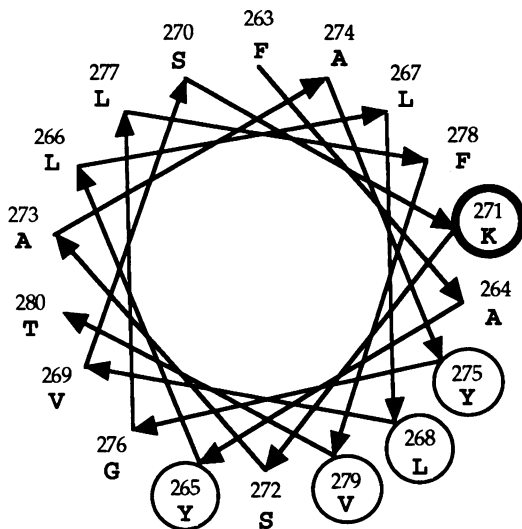


Fig. 5. Helical wheel representation of the axolotl TCR $C\beta 1$ TM sequence from Phe-263 to Thr-280. Conserved residues (circled) are clustered on one face of the helix; the lysine (K) residue is circled with a thick line.

the organization of the TCR $C\beta$ EC region into an immunoglobulin-like domain consisting of multistrand β -sheet bilayers (Fig. 4B). The extra cysteine (Cys-191) residue that is present in all known mammalian $C\beta$ chains is lacking in both axolotl and chicken $C\beta$ chains (3). The hinge region appears to be well conserved, including the Cys-247 residue, which is thought to covalently interact with the TCR α chains. A majority of the hydrophobic amino acids are found in the 36 residues of the transmembrane (TM) region. However, four charged hydrophobic residues (Asp-252, Arg-255, Lys- or Asn-259, and Arg-262) are present at the N-terminal end of the TM region and may not in fact be buried in the membrane bilayer. The TM region carries a positively charged lysine (Lys-271) flanked by strongly hydrophobic tyrosine, leucine, and valine residues. These amino acids would all lie on one side of an α -helix if the TM segment adopted such a structure (Fig. 5) and may form an important surface interaction with another TM domain belonging to a putative polypeptide that could be an element of a conserved CD3-like molecule (28).

In Table 1, the axolotl $C\beta 1$ and its putative EC and TM domains are compared to the corresponding axolotl $C\beta 2$, chicken $C\beta$ (3), and human $C\beta 2$ (18) regions. At the nucleotide level, the human and chicken EC domains present >55% identical residues with the corresponding axolotl regions. The axolotl $C\beta 1$ and $C\beta 2$ TM domains are equally divergent between themselves and with the chicken TM domain but are less similar to the human TM domain.

Our results show that the overall structure of the TCR β chain is conserved in tetrapods and that it contains elements that in mammalian species are involved in molecular interactions with the companion TCR α chain and some of the

Table 1. Percent of nucleotide and amino acid matches for the axolotl $C\beta 2$, chicken $C\beta$, and human $C\beta 2$ chains relative to the axolotl $C\beta 1$ chain

| Chain | % identical residues | | |
|--------------------|--------------------------|----------------------|------------------------|
| | Overall $C\beta 1$ chain | EC domain (aa 1–111) | TM domain (aa 118–154) |
| Axolotl $C\beta 2$ | 85.9/80.5 | 94.7/90.0 | 60.0/63.8 |
| Chicken $C\beta$ | 54.0/32.8 | 55.6/43.5 | 59.0/41.6 |
| Human $C\beta 2$ | 51.4/41.0 | 56.0/45.4 | 37.2/25.0 |

The first number represents percent identities at the nucleotide level and the second number is at the amino acid level.

polypeptides that form the CD3 complex. This observation is reinforced by the recent characterization of a surface multi-molecular complex on the immunoglobulin-negative axolotl peripheral lymphocytes that could be the amphibian equivalent of the advanced vertebrate CD3/TCR complex (F.K., F.G., J.C., and A. Tournefier, unpublished data). The very large size of the axolotl genome [≈ 42.0 pg of DNA per haploid genome (29)] led to considerable theoretical and technical difficulties for the construction and screening of representative genomic libraries and for the determination of restriction maps by the Southern blot hybridization technique. Thus it is not actually possible to give an estimation of the potential number of the $C\beta$ genomic elements or to estimate the number of $V\beta$ genes correlated to the $V\beta 22^+(1)$ segment.

We thank Joseph Schwager, Michael V. Wiles, and Patrice Marche for help and valuable discussion in several phases of this work and Gillian Butler-Browne for assistance in preparing the manuscript. This work was supported by the Université Pierre et Marie Curie and the Centre National de la Recherche Scientifique (Unité de Recherche Associée 1135).

- Allison, J. P. & Lanier, L. L. (1987) *Annu. Rev. Immunol.* **5**, 503–540.
- Brenner, M. B., Strominger, J. L. & Krangel, M. S. (1988) *Adv. Immunol.* **43**, 133–191.
- Tjoelker, L. W., Carlson, L. M., Lee, K., Lahti, J., McCormack, W. T., Lieden, J. M., Chen, C.-L., Cooper, M. D. & Thompson, C. B. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7856–7860.
- Davis, M. M. (1990) *Annu. Rev. Biochem.* **59**, 475–496.
- Du Pasquier, L. (1989) in *Fundamental Immunology*, ed. Paul, E. (Raven, New York), 2nd Ed., pp. 139–165.
- Williams, A. F. & Barclay, A. N. (1988) *Annu. Rev. Immunol.* **6**, 381–405.
- Shablott, M. J. & Litman, G. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4684–4688.
- Fellah, J. S. & Charlemagne, J. (1988) *Mol. Immunol.* **25**, 1377–1386.
- Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
- Thiesen, H. J., Casorati, G., Lauster, R. & Wiles, M. V. (1991) in *Immunological Methods*, eds. Lefkowitz, I. & Pernis, B. (Academic, London), Vol. 4, pp. 35–50.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5464–5467.
- Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, G. (1991) *Sequences of Proteins of Immunological Interest* (Natl. Inst. Health, Bethesda, MD), 5th Ed.
- Fellah, J. S., Wiles, M. V., Charlemagne, J. & Schwager, J. (1992) *Eur. J. Immunol.* **22**, 2595–2601.
- Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2449.
- Johnson, N. A., Carland, F., Allen, P. M. & Glimcher, L. H. (1989) *J. Immunol.* **142**, 3298–3304.
- Williams, C. B. & Gutman, G. A. (1989) *J. Immunol.* **142**, 1027–1035.
- Wilson, R. K., Lai, E., Concannon, P., Barth, R. K. & Hood, L. Y. (1988) *Immunol. Rev.* **101**, 149–172.
- Tunacliffe, A., Kefford, R., Milstein, C., Forster, A. & Rabbitts, T. H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5068–5072.
- Schiffer, M., Wu, T. T. & Kabat, E. A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4461–4463.
- Milner, A. R. (1988) in *The Phylogeny and Classification of Tetrapods*, ed. Benton, M. J. (Clarendon, Oxford), Vol. 1, pp. 59–102.
- Jouvin-Marche, E., Heller, M. & Rudikoff, S. J. (1986) *J. Exp. Med.* **164**, 2083–2088.
- Clarke, S. H. & Rudikoff, S. (1984) *J. Exp. Med.* **159**, 773–782.
- Rudikoff, S., Fitch, W. M. & Heller, M. (1992) *Mol. Biol. Evol.* **9**, 14–26.
- Tanaka, A., Ishiguro, N. & Shinagawa, M. (1990) *Immunogenetics* **32**, 263–271.
- Angioliello, A. L., Lamoyi, E., Bernstein, K. E. & Mage, R. G. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8624–8628.
- Gascoigne, N. R. J., Chien, Y.-H., Becker, D. M., Kavalier, J. & Davis, M. M. (1984) *Nature (London)* **310**, 387–391.
- Novotny, J., Tonegawa, S., Saito, H., Kranz, D. M. & Eisen, H. N. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 742–746.
- Alcover, A., Marinza, R. A., Ermonval, M. & Acuto, O. (1990) *J. Biol. Chem.* **265**, 4131–4135.
- Britten, R. J. & Davidson, E. H. (1971) *Quart. Rev. Biol.* **46**, 111–133.