



Published in final edited form as:

*Immunogenetics*. 2005 March ; 56(12): 924–929. doi:10.1007/s00251-004-0766-y.

## Variable domains and a VpreB-like molecule are present in a jawless vertebrate

### John P. Cannon,

Department of Molecular Genetics, All Children's Hospital, 801 Sixth Street South, St. Petersburg, FL, 33701, USA, Tel.: +1-727-553-3601, Fax: +1-727-553-3610; H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Avenue, Tampa, FL, 33612, USA, Department of Pediatrics, USF/ACH Children's Research, Institute, University of South Florida College of Medicine, 830 First Street South, St. Petersburg, FL, 33701, USA

### Robert N. Haire,

Department of Pediatrics, USF/ACH Children's Research, Institute, University of South Florida College of Medicine, 830 First Street South, St. Petersburg, FL, 33701, USA

### Zeev Pancer,

Howard Hughes Medical Institute, University of Alabama at Birmingham, 378 Wallace Tumor Institute, Birmingham, AL, 35294, USA

### M. Gail Mueller,

Department of Pediatrics, USF/ACH Children's Research, Institute, University of South Florida College of Medicine, 830 First Street South, St. Petersburg, FL, 33701, USA

### Diana Skapura,

Department of Pediatrics, USF/ACH Children's Research, Institute, University of South Florida College of Medicine, 830 First Street South, St. Petersburg, FL, 33701, USA

### Max D. Cooper, and

Howard Hughes Medical Institute, University of Alabama at Birmingham, 378 Wallace Tumor Institute, Birmingham, AL, 35294, USA

### Gary W. Litman

Department of Molecular Genetics, All Children's Hospital, 801 Sixth Street South, St. Petersburg, FL, 33701, USA, Tel.: +1-727-553-3601, Fax: +1-727-553-3610; H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Avenue, Tampa, FL, 33612, USA; Department of Pediatrics, USF/ACH Children's Research, Institute, University of South Florida College of Medicine, 830 First Street South, St. Petersburg, FL, 33701, USA

John P. Cannon: litmang@allkids.org; Gary W. Litman: litmang@allkids.org

## Abstract

Immunoglobulins (Igs) and T cell antigen receptors (TCRs) that undergo somatic diversification have not been identified in the two extant orders of jawless vertebrates, which occupy essential positions in terms of understanding the evolution of the emergence of adaptive immunity. Using a single motif-dependent PCR-based approach coupled with a vector that allows selection of cDNAs encoding secretion signal sequences, four different genes encoding Ig V-type domains were identified in the sea lamprey (*Petromyzon marinus*). One of the predicted proteins encoded by these genes shares structural characteristics with mammalian VpreB molecules, including the

absence of a recognizable transmembrane region, a relatively high proportion of charged amino acids in its C-terminal tail and distinctive features of its secretion signal peptide. This is the first indication of a molecule related to the B cell receptor (BCR) complex in a species that diverged prior to the jawed vertebrates in which RAG-mediated adaptive immunity is first encountered.

## Keywords

Jawless vertebrate; Lymphocyte ontogeny; V-related sequence; Phylogeny; Immunoglobulin; T cell antigen receptor

## Introduction

The adaptive immune response is mediated by segmental genes encoding immunoglobulins (Igs) and T cell antigen receptors (TCRs), which undergo somatic reorganization in B and T lymphocytes, respectively. Ig and TCR, as well as major histocompatibility complex (MHC) class I and MHC class II, are present in all species of jawed vertebrates examined to date, as are lymphocytes and lymphoid organs resembling those found in mammals (Litman et al. 1999). Furthermore, many aspects of the transcriptional regulation of lymphoid development in even the most phylogenetically primitive jawed vertebrates are similar to those found in mammals (Anderson et al. 2001; Cupit et al. 2003; Haire et al. 2000).

Jawless vertebrates, of which lampreys and hagfish are the sole extant orders, are the most phylogenetically advanced metazoans that do not possess classical RAG-mediated immunity, making further understanding of immune phenomena in these species critical to our interpretation of immune system evolution. Efforts to identify conventional Ig, TCR, MHC I or MHC II in these species have been unsuccessful using both conventional hybridization strategies and extensive expressed sequence tag (EST) screens (Mayer et al. 2002; Uinuk-Ool et al. 2002). However, the sea lamprey (*Petromyzon marinus*) exhibits allograft rejection and responds to immunization by production of humoral molecules that lack the common features of Igs found throughout the jawed vertebrates (Litman et al. 1970, 1999). Furthermore, cells resembling lymphocytes have been described in the sea lamprey (Finstad et al. 1969; Mayer et al. 2002) and it has been shown recently that these cells transcribe a novel variable lymphocyte receptor (VLR) consisting of leucine-rich repeats (LRRs) that generate diversity through a somatic rearrangement process (Pancer et al. 2004a). The relationship of this molecule to those previously shown to be associated with induced immunity to bacteria and human erythrocytes is presently unclear.

Two reports of EST sequencing in lamprey failed to identify gene products that encode an Ig-type variable (V) region, a central feature of the rearranging antigen binding receptors. On the other hand, IgV-type domains have been described in several mammalian proteins that are not necessarily associated with antigen binding, as well as in pro-tochordates (Cannon et al. 2002; Sato et al. 2003). Given the inherent limitations of EST screening procedures, the possibility exists that structural homologs of a B cell receptor (BCR) or TCR, or molecules related to the expression or function of such structures, could be present in this species. In order to explore this possibility, we have utilized a previously described cloning approach that requires knowledge (or prediction) of only a single short amino acid sequence motif from secreted or membrane proteins (Cannon et al. 2002) and herein describe several different genes possessing V-type regions in the sea lamprey. One of these genes is predicted to encode a secreted molecule that consists of a single Ig V domain and shares structural features with VpreB, an Ig surrogate light chain component expressed early in mammalian B cell ontogeny and integral to BCR presentation (Karasuyama et al. 1996; Kudo and Melchers 1987). Although the function of the product of this gene is not

understood, these studies raise the possibility that the cell type designated as a lymphocyte in this species also expresses a molecule integral to the genesis of a BCR.

## Materials and methods

### Amptrap PCR and cloning of cDNA from sea lamprey tissues

Amptrap PCR was performed as described (Cannon et al. 2002). Briefly, mRNA from sea lamprey tissues (adult peripheral blood leukocytes, larval opisthonephros and typhlosole) was reverse transcribed using the SMART cDNA procedure (Clontech, Palo Alto, Calif). 5'-RACE PCR was performed using a SMART oligonucleotide combined with *Sfi*I-flanked degenerate oligonucleotides complementing codons for the following amino acid sequences (provided in single-letter abbreviation): CHVEH, C(H/Q)V (D/E)H, WFK, WYX or WX, where "X" designates any amino acid. After *Sfi*I digestion of the PCR products, ligation into the *Sfi*I-linearized Amptrap vector G7311, and transformation of *Escherichia coli* DH10 $\beta$ , transformants bearing signal peptide-encoding plasmids were selected on LB plates containing ampicillin and kanamycin. Insert size was surveyed using vector-specific primers and direct colony PCR; colony PCR amplicons were sequenced directly using dye primer chemistry on a LI-COR IR4200 DNA sequencer.

### cDNA Library

A lamprey leukocyte cDNA library was constructed in  $\lambda$ TriplEx2 using the SMART library protocol (Clontech). This library was screened successfully for full-length cDNAs of PM3033 and PM7055 using hybridization probes derived from the initial Amptrap clones. To extend the coding regions of PM3601 and PM6889, 3'-RACE PCR was performed on library aliquots using appropriate combinations of gene- and vector-specific primers.

### Virtual northern blot

Larval lamprey lymphocytes were stimulated using a cocktail of bacteria, erythrocytes, and mitogens (at weekly intervals for approximately three weeks) and subsequently sorted for lymphocyte or myeloid morphology by FACS (Pancer et al. 2004a). Lymphocytes from unstimulated (naïve) larvae also were isolated for comparison to stimulated cells. cDNA samples from each population or, alternatively, from various lamprey tissues, were synthesized using the Clontech SuperSMART technique. Each cDNA sample was amplified using the 5' and 3' anchor-specific PCR primers. Amplified cDNA samples were separated by agarose gel electrophoresis and transferred to a charged nylon membrane for hybridization with various radiolabeled probes that included the coding sequences of PM3033, PM3601, and *P. marinus* glyceraldehyde 3-phosphate dehydrogenase (GAPDH).

## Results

We previously described Amptrap 5'-RACE PCR, which allows selection for cloned cDNAs encoding secretion signal sequences by virtue of complementation of a signal peptide-defective  $\beta$ -lactamase gene in *E. coli* (Cannon et al. 2002). Application of Amptrap 5'-RACE cloning from different sea lamprey cDNA templates (adult white blood cell, larval opisthonephros or larval typhlosole) using various primers yielded four amplicons possessing structural characteristics of typical Ig V-type domains; 3'-RACE and cDNA library screening were used to extend the cDNA sequences. The V domains encoded by each cDNA are aligned in Fig. 1. Based on BLAST analysis, *PM6889* is a candidate homolog of mammalian nectin, a transmembrane molecule with multiple isoforms. Of note, *PM6889* was identified using an Amptrap PCR primer corresponding to only a single defined amino acid (Trp-X). *PM7055* is a single Ig domain transmembrane protein with an 81 amino acid cytoplasmic tail. The Ig domain is a long (~120 amino acid) V-type sequence with canonical

V-frame residues (Harpaz and Chothia 1994) and additional sequence similarity to some Igs and TCRs at the predicted peptide level. The product of the partial cDNA *PM3601* is a 221 amino acid protein homologous to the Lutheran blood group antigen (LuBGA) and gicerin, both of which are cell surface adhesion molecules containing a V-type domain. A *PM3601* exon/intron boundary in a genomic clone recovered from a *Petromyzon* phage  $\lambda$  library is identical to the respective boundary for the equivalent exon in human LuBGA and in human gicerin (data not shown). The fourth gene, *PM3033*, amplified from both adult white blood cells and larval typhlosole (in which lymphocytes are present), encodes a 156 amino acid polypeptide that preserves all of the canonical V-frame residues in its single V-type domain (Cannon et al. 2002; Harpaz and Chothia 1994). Unlike the other three genes, which encode transmembrane proteins, *PM3033* encodes a putative secreted molecule with an elevated content of positive- and negative-charged amino acids in a uniquely extended region C-terminal to the V domain. Sites for glycosylphosphatidylinositol (GPI)-linked membrane anchoring (Eisenhaber et al. 1999) are not apparent in the C-terminal sequence.

To our knowledge, single V domain proteins that are secreted and not GPI-linked are limited essentially to mammalian VpreB, which is expressed early in B cell ontogeny and is distinguished from the V regions of rearranging antigen binding receptors by the absence of a J sequence (Kudo and Melchers 1987). VpreB possesses a C-terminal region with a high density of charged amino acids, as also is seen in *PM3033*. Furthermore, like most mammalian VpreBs (and closely related genes), *PM3033* encodes dual cysteine residues in its leader, as does the lamprey gene *PM3601* described here, as well as a proline near the site of signal peptide cleavage, a distinguishing characteristic of both human and mouse VpreB (Minegishi et al. 1999). A WE amino acid motif is present in the C-terminal domains of all three proteins. BLAST analyses of the V domain of *PM 3033* indicate modest identity with the V domains of several known molecules, including epithelial V-like antigen (23% amino acid identity) (Guttinger et al. 1998), numerous human Ig V<sub>H</sub> regions (approximately 23% amino acid identity), and *Branchiostoma* (amphioxus) VDB (26% amino acid identity) (Sato et al. 2003) (Fig. 2).

Recently it was shown that lymphocyte-like cells in sea lamprey can be activated in response to stimulation with a cocktail of bacteria, sheep erythrocytes and plant mitogens (Pancer et al. 2004a). In order to examine the possibility that the expression of *PM3033* is altered in sea lamprey leukocyte populations after activation, a “virtual” Northern blot of cDNA from FACS-separated activated lamprey leukocytes and other lamprey tissues was hybridized with a probe complementing *PM3033*. Preferential hybridization of a *PM3033* cDNA probe was observed in lamprey leukocytes (Fig. 3). Transcripts in large activated blood lymphocytes and a population of myeloid cells that contains activated lymphocytes (~5%) appear to express the gene at higher levels relative to unstimulated leukocytes and activated small lymphocyte-like cells. In contrast, *PM3033* transcripts are not detected in erythrocyte, liver or tail samples.

## Discussion

Defining the immune systems of the jawless vertebrates is critical to our understanding of the origins of adaptive immunity. It is important to recognize in this regard that inferences relating to the absence of conventional immune receptors in these species are based on negative findings, although the cross hybridization and PCR-based identification strategies on which these conclusions have been drawn have successfully identified Ig, TCR, MHC I, MHC II, terminal deoxynucleotidyltransferase and recombination activating genes in the most phylogenetically ancient of the extant lineages of jawed vertebrates (Litman et al. 1999). EST-based approaches offer greater possibilities for identifying distantly related homologous genes. Large-scale EST surveys conducted in lamprey identified several

candidate orthologs of genes that are expressed in mammalian lymphocytes whose expression is not solely lymphocyte-specific. Neither V domain-containing nor other Ig-type genes were identified in these studies (Mayer et al. 2002; Uinuk-Ool et al. 2002).

Although mRNA (cDNA) representation can be normalized, EST surveys can be biased in terms of identification of rare and/or developmentally regulated mRNAs. As demonstrated by the cloning of the nectin-like gene *PM6889* in this study, Amtrap selection can identify Ig-superfamily cDNAs based on the presence of only minimal amino acid motifs, e.g., Trp-X. Although such short motifs are essentially ubiquitous among proteins, the Amtrap method filters only those cDNAs encoding secretion signal sequences and of a specified length from a given tissue. Of the two primary strategies employed here, the one directed at the highly conserved tryptophan residue in the V region has been the more productive (J. Cannon, unpublished observations) and provides one explanation (see below) for why a recently identified lamprey gene with certain TCRV features (Pancer et al. 2004b) was not amplified as it lacks this residue.

Diversified families of immune-type receptors are found in both invertebrates (Adema et al. 1997; Pancer 2000) and protochordates (Cannon et al. 2002); in the case of the former, the immune-protective molecule undergoes genomic and somatic change that is associated with specific recognition of a parasite (Zhang et al. 2004). In jawless vertebrates, the VLR genes encode a diverse set of proteins containing variable N-terminal LRR motifs and a constant C-terminus. Extensive variation in VLR transcripts is created by individual lymphocyte-restricted somatic rearrangement of germline LRR cassettes, resulting in a unique VLR transcript in each cell. The VLR ectodomain is GPI-anchored to the plasma membrane and can be liberated from the surface of the individual cells (Pancer et al. 2004a). Although specific antigen recognition by VLRs has yet to be reported, these molecules likely play a significant role in the sea lamprey immune response. Therefore, in a general sense, it is now necessary to consider that the adaptive immune response could be mediated through at least two unrelated mechanisms. In jawed vertebrates, such responses involve a RAG-mediated recombination process involving members of the IgSF. In jawless vertebrates, adaptive immunity does not appear to be RAG-mediated, likely does not exclusively involve members of the Ig superfamily and could involve more than one family of receptors. It is not unreasonable to assume that yet other mechanisms that create diversity in immune-type receptors exist in invertebrates and protochordates, as suggested by recent reports (Cannon et al. 2004; Zhang et al. 2004).

The findings reported here define a gene product that suggests a relationship between molecules that are associated with B cell immunity in jawed and now jawless vertebrates. Specifically, the predicted sequence of *PM3033* shares features with VpreB, including similar short sequence motifs, predicted mass, isoelectric point and net C-terminal region charge. The modest relatedness (23% absolute homology; 40–42% net similarity) of the core V region of *PM3033* to the V regions of human and mouse VpreB is not surprising given the large phylogenetic distance separating the jawless and jawed vertebrate forms. Furthermore, the VpreB molecules found in mammals and likely other jawed vertebrates interact with specific structural features of the mammalian Ig heavy chain, as well as with  $\lambda 5$ , the second component of the surrogate light chain (Bauer et al. 1988; Karasuyama et al. 1990; Lassoued et al. 1993). Notably, we have been unable to identify lamprey Ig heavy chain homologs using either low-stringency DNA hybridization, which can detect less than 60% absolute nucleotide sequence identity, or short-primer PCR, which relies on evolutionary conservation of a pair of amino acid motifs 3–4 residues in length. Thus, if *PM3033* is a distant ortholog of VpreB and interacts with a structure equivalent to an Ig heavy chain in the sea lamprey, sequence conservation between *PM3033* and mammalian VpreB is predicted to be remote. Efforts to localize VpreB binding to the lymphocyte cell surface

using polyclonal antibodies to *PM3033* have so far been inconclusive but identification of mammalian VpreB by conventional cell surface immunofluorescence also is very difficult because of low expression levels (Wang et al. 2002). Nevertheless, expression of the *PM3033* gene in immunologically stimulated, sorted larval lymphocytes, and possibly myeloid cells, appears to be increased relative to unstimulated control populations. Whether *PM3033* expression reflects upregulation in the stimulated populations or, alternatively, represents a shift in relative proportions of PM3033-expressing cell types after activation currently is unknown.

The recent description of a non-rearranging single copy gene sequence in lamprey that can be modeled to a TCR V suggests that other molecules that are related to the combinatorial antigen binding receptors may exist in jawless vertebrates (Pancer et al. 2004b). Further characterization of the *PM3033* gene product, particularly its ontogenetic expression and association with other cell surface molecules, could address the possibility that a molecule that plays no direct role in antigen binding but has a significant place in B cell ontogeny arose independently from the antigen-binding receptors seen in contemporary jawed vertebrates. As such, further studies of its function could provide new insight into the origins of the modern forms of BCRs.

## Acknowledgments

We thank Ronda Litman for DNA sequence analysis and Barbara Pryor for editorial assistance. This work was supported by grants to G.W.L. and M.D.C. from the National Institutes of Health, and by a grant from the National Science Foundation to Z.P. M.D.C. is an Investigator of the Howard Hughes Medical Institute and Z.P. was a recipient of the Cottrell postdoctoral award

## References

- Adema CM, Hertel LA, Miller RD, Loker ES. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. *Proc Natl Acad Sci USA*. 1997; 94:8691–8696. [PubMed: 9238039]
- Anderson MK, Sun X, Miracle AL, Litman GW, Rothenberg EV. Evolution of hematopoiesis: three members of the PU.1 transcription factor family in a cartilaginous fish, *Raja eglanteria*. *Proc Natl Acad Sci USA*. 2001; 98:553–558. [PubMed: 11149949]
- Bauer SR, Kudo A, Melchers F. Structure and pre-B lymphocyte restricted expression of the *VpreB* gene in humans and conservation of its structure in other mammalian species. *EMBO J*. 1988; 7:111–116. [PubMed: 3258819]
- Cannon JP, Haire RN, Litman GW. Identification of diversified immunoglobulin-like variable region-containing genes in a protochordate. *Nature Immunol*. 2002; 3:1200–1207. [PubMed: 12415263]
- Cannon JP, Haire RN, Schnitker N, Mueller MG, Litman GW. Individual protochordates possess unique immune-type receptor repertoires. *Curr Biol*. 2004; 14:R465–R466. [PubMed: 15203016]
- Cupit PM, Hansen JD, McCarty AS, White G, Chioda M, Spada F, Smale ST, Cunningham C. Ikaros family members from the agnathan *Myxine glutinosa* and the urochordate *Oikopleura dioica*: emergence of an essential transcription factor for adaptive immunity. *J Immunol*. 2003; 171:6006–6013. [PubMed: 14634112]
- Eisenhaber B, Bork P, Eisenhaber F. Prediction of potential GPI-modification sites in proprotein sequences. *J Mol Biol*. 1999; 292:741–758. [PubMed: 10497036]
- Finstad, J.; Fänge, R.; Good, RA. The development of lymphoid systems: immune response and radiation sensitivity in lower vertebrates. In: Fiore-Donati, L.; Hanna, MG., editors. *Lymphatic tissue and germinal centers in immune response*. Plenum; New York: 1969. p. 21-31.
- Guttinger M, Sutti F, Panigada M, Porcellini S, Merati B, Mariani M, Teesalu T, Cons GG, Grassi F. Epithelial V-like antigen (EVA), a novel member of the immunoglobulin superfamily, expressed in embryonic epithelia with a potential role as homotypic adhesion molecule in thymus histogenesis. *J Cell Biol*. 1998; 141:1061–1071. [PubMed: 9585423]

- Haire RN, Miracle AL, Rast JP, Litman GW. Members of the Ikaros gene family are present in early representative vertebrates. *J Immunol.* 2000; 165:306–312. [PubMed: 10861066]
- Harpaz Y, Chothia C. Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J Mol Biol.* 1994; 238:528–539. [PubMed: 8176743]
- Karasuyama H, Kudo A, Melchers F. The proteins encoded by the *VpreB* and *lambda 5* pre-B cell-specific genes can associate with each other and with mu heavy chain. *J Exp Med.* 1990; 172:969–972. [PubMed: 2117638]
- Karasuyama H, Rolink A, Melchers F. Surrogate light chain in B cell development. *Adv Immunol.* 1996; 47:1–41. [PubMed: 8787628]
- Kudo A, Melchers F. A second gene, *VpreB* in the *I5* locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes. *EMBO J.* 1987; 6:2267–2272. [PubMed: 3117530]
- Lassoued K, Nunez CA, Billips L, Kubagawa H, Monteiro RC, LeBlen TW, Cooper MD. Expression of surrogate light chain receptors is restricted to a late stage in pre-B cell differentiation. *Cell.* 1993; 73:73–86. [PubMed: 7681728]
- Litman GW, Frommel D, Finstad J, Howell J, Pollara BW, Good RA. The evolution of the immune response. VIII. Structural studies of the lamprey immunoglobulin. *J Immunol.* 1970; 105:1278–1285. [PubMed: 4099053]
- Litman GW, Anderson MK, Rast JP. Evolution of antigen binding receptors. *Annu Rev Immunol.* 1999; 17:109–147. [PubMed: 10358755]
- Mayer WE, Uinuk-Ool T, Tichy H, Gartland LA, Klein J, Cooper MD. Isolation and characterization of lymphocyte-like cells from a lamprey. *Proc Natl Acad Sci USA.* 2002; 99:14350–14355. [PubMed: 12388781]
- Minegishi Y, Hendershot LM, Conley ME. Novel mechanisms control the folding and assembly of 15/14.1 and VpreB to produce an intact surrogate light chain. *Proc Natl Acad Sci USA.* 1999; 96:3041–3046. [PubMed: 10077633]
- Pancer Z. Dynamic expression of multiple scavenger receptor cysteine-rich genes in coelomocytes of the purple sea urchin. *Proc Natl Acad Sci USA.* 2000; 97:13156–13161. [PubMed: 11069281]
- Pancer Z, Amemiya CT, Ehrhardt RA, Ceitlin J, Gartland GL, Cooper MD. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature.* 2004a; 430:174–180. [PubMed: 15241406]
- Pancer Z, Mayer WE, Klein J, Cooper MD. Prototypic T cell receptor and CD4-like coreceptor are expressed by lymphocytes in the agnathan sea lamprey. *Proc Natl Acad Sci USA.* 2004b; 101:13273–13278. [PubMed: 15328402]
- Sato A, Mayer WE, Klein J. A molecule bearing an immunoglobulin-like V region of the CTX subfamily in amphioxus. *Immunogenetics.* 2003; 55:423–427. [PubMed: 12898067]
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994; 22:4673–4680. [PubMed: 7984417]
- Uinuk-Ool T, Mayer WE, Sato A, Dongak R, Cooper MD, Klein J. Lamprey lymphocyte-like cells express homologs of genes involved in immunologically relevant activities of mammalian lymphocytes. *Proc Natl Acad Sci USA.* 2002; 99:14356–14361. [PubMed: 12391333]
- Wang YH, Stephan RP, Scheffold A, Kunkel D, Karasuyama H, Radbruch A, Cooper MD. Differential surrogate light chain expression governs B-cell differentiation. *Blood.* 2002; 99:2459–2467. [PubMed: 11895780]
- Zhang SM, Adema CM, Kepler TB, Loker ES. Diversification of Ig superfamily genes in an invertebrate. *Science.* 2004; 305:251–254. [PubMed: 15247481]

```

PM3033  MVISFFILCLVCVHPLEAVSVTLMPHGSVS--AGKVLTLTCNFT----SDAKTSPQTWMT
PM3601  MNTLCVTCILLTTLPGVITLLSVKTPTELTVITGSDLEIPCQWNVPQGDGKSSSDIILQ
PM6889  MIAVQLLLLLSATLGAQSVTSKIDHDSEASGSGGDAELRCVY-----KGSKTARQIIWE
PM7055  -MHGFCKCVVLPLCVLIVLSALDAHNNMAVFVGQNVSLQCPTVS---RVSPKQPTVHWE

PM3033  WSFGNAEVCKYVNNNLRPISNEAFKH-----RTCHGDIKRN-----FIQMSIS
PM3601  WFIETDKGERRRVFYKKGHLQVREKP-----EGYAREMLVGK-----DFTMLLK
PM6889  KETNQERVIASRHDDKSHVSDVTLQ-----ERLTFLGQGNN-----DATILIP
PM7055  WTGKNHRAPLQLAEGSLSHGTFNARYNIDALRMFGRRLALWRQRGSPSWRQDVEFTLGIA

PM3033  NIQVNDSGIYKC
PM3601  NVTRQDSRGYIC
PM6889  PLRVSDSGTYIC
PM7055  HVRSSDQGVFTC

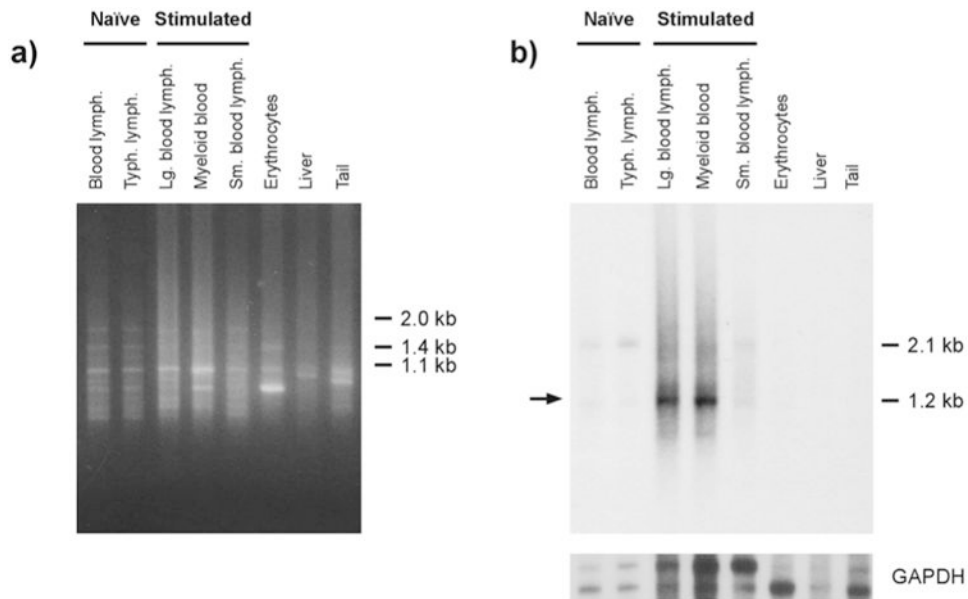
```

**Fig. 1.** ClustalW alignment (Thompson et al. 1994) of four predicted immunoglobulin variable (Ig V)-type regions from the sea lamprey (*Petromyzon marinus*). Each sequence shown extends from the initiation methionine to the second cysteine of the canonical V-frame. Predicted secretion signal peptides are *italicized*. V-frame residues are shaded in *black*





**Fig. 2.** ClustalW alignment of the predicted protein product of *PM3033* with human and mouse VpreB amino acid sequences [GenBank accession nos. NP\_009059 (human) and CAA29072 (mouse)]. Identical amino acids shared by all three proteins are shaded in *black*; biochemically similar amino acids shared by all three proteins are shaded in *gray*. Charged amino acids at the C-termini of the proteins are indicated by a *boxed outline*. Stop codons are indicated by “.”



**Fig. 3.** Virtual Northern cDNA blotting assay indicating expression of *PM3033* transcripts in sea lamprey larvae. **a** Ethidium bromide staining of cDNA samples after 1% agarose gel electrophoresis. **b** Autoradiogram after hybridization of *PM3033* (*top*) and glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*, *bottom*) cDNA sequences to the blotted cDNA from (**a**). Parallel hybridization with a probe complementing *PM3601* failed to demonstrate this effect (data not shown). Identities of samples are indicated at the *top*. *Typh* Typhlosole, *lymph* lymphocyte-like cells, *Lg* large, *Sm* small, the *arrow* indicates the relative position of the *PM3033* transcript. Approximate molecular sizes are indicated to the *right* of each image. The *GAPDH* probe hybridizes to two major transcripts, approximately 1.5 and 1.7 kilobases (kb) in length