Version of Record: https://www.sciencedirect.com/science/article/pii/S0016648016300028
Manuscript_762463aa7eaceaabe89fe17c0e27db64

The emergence of the vasopressin and oxytocin hormone receptor gene family lineage: clues from the characterization of vasotocin receptors in the sea lamprey (*Petromyzon marinus*)

Sally A. Mayasich* and Benjamin L. Clarke

Department of Biomedical Sciences
School of Medicine
University of Minnesota-Duluth
1035 University Drive
Duluth, MN 55812
USA

*Corresponding author. Email: maya0010@d.umn.edu; Phone: 218-428-1531
Senior author email: bclarke@d.umn.edu

© 2016 published by Elsevier. This manuscript is made available under the Elsevier user license
https://www.elsevier.com/open-access/userlicense/1.0/

Abstract

The sea lamprey (*Petromyzon marinus*) is a jawless vertebrate at an evolutionary nexus between invertebrates and jawed vertebrates. Lampreys are known to possess the arginine vasotocin (AVT) hormone utilized by all non-mammalian vertebrates. We postulated that the lamprey would possess AVT receptor orthologs of predecessors to the arginine vasopressin (AVP)/oxytocin (OXT) family of G protein-coupled receptors found in mammals, providing insights into the origins of the mammalian V1A, V1B, V2 and OXT receptors. Among the earliest animals to diverge from the vertebrate lineage in which these receptors are characterized is the jawed, cartilaginous elephant shark, which has genes orthologous to all four mammalian receptor types. Therefore, our work was aimed at helping resolve the critical gap concerning the outcomes of hypothesized large-scale (whole-genome) duplication events. We sequenced one partial and four full-length putative lamprey AVT receptor genes and determined their mRNA expression patterns in 15 distinct tissues. Three of the full-coding genes possess structural characteristics of the V1 clade containing the V1A, V1B and OXT receptors. Another full-length coding gene and the partial sequence are part of the V2 clade and appear to be most closely related to the newly established V2B and V2C receptor subtypes. Our synteny analysis also utilizing the Japanese lamprey (*Lethenteron japonicum*) genome supports the recent proposal that jawless and jawed vertebrates shared one-round (1R) of WGD as the most likely scenario.

Keywords:

Vasotocin; G protein-coupled receptor; Agnathan vertebrate; Sea lamprey; Molecular evolution; Whole-genome duplication

1. Introduction

The sea lamprey (*Petromyzon marinus*) is a jawless (agnathan) vertebrate species whose ancestors diverged from the vertebrate lineage approximately 500 million years ago. Lampreys have become key organisms in studying the origin of jawed vertebrates from jawless ancestors. As demonstrated in studies of gonadotropin releasing hormone (GnRH) (Sower et al., 2012), melanocortin (Haitina et al., 2007) and growth hormone (GH)/prolactin (PRL) (Ellens et al., 2013) receptor families, the characterization of neuroendocrine hormone receptor groups in the sea lamprey has immense value in evolutionary studies. These findings are important in deciphering the time sequence of duplication and divergence within gene family lineages, as well as viewing the larger picture of whole-genome duplications (WGD) and speciation. Two rounds of WGD (1R and 2R) early in the vertebrate lineage have been proposed (Ohno, 1970) and further supported by findings on the emergence of chordates (Putnam et al., 2008), but the timing of the events with respect to cyclostome-gnathostome divergence is an ongoing topic of study. A major question has been whether the agnathans diverged from the vertebrate lineage prior to 1R, between 1R and 2R, or after 2R (Kuraku et al., 2009; Smith et al., 2013). It has been proposed that a third round (3R) of WGD took place in the lamprey lineage after divergence, or possibly all three rounds occurred independently of the gnathostome 2R WGDs (Mehta et al., 2013; and Nah et al., 2014). The most recent studies, however, contest the 2R WGD scenario (Smith and Keinath, 2015). A lamprey meiotic linkage map has been constructed to provide near chromosome-level evidence supporting an agnathan-gnathostome shared 1R WGD with subsequent segmental duplications, fusion and translocation, rather than a second WGD in either lineage (Smith and Keinath, 2015).

The arginine vasopressin (AVP) and oxytocin (OXT) hormone family, including vasotocin (AVT), and the corresponding G protein-coupled receptor (GPCR) family have been

characterized in species representing the vertebrate classes Chondrichthyes, Osteichthyes, Amphibia, Reptilia, Aves and Mammalia. Although the AVT preprohormone has been sequenced from the Arctic (Japanese) lamprey (*Lethenteron camtschaticum*, synonym *Lethenteron japonicum*) (Gwee et al., 2009), there have been no studies characterizing the AVT receptors in these agnathans. This constitutes a gap in our evolutionary knowledge of an important hormone-receptor system that could aid in resolving the WGD timing issue.

Gwee et al. (2009) provide evidence that after agnathan divergence from the vertebrate lineage, a tandem duplication of a neurohypophysial hormone occurred in a common ancestor of jawed vertebrates resulting in two hormone paralogs. Thus, only one hormone, AVT, has been found in agnathans. In most mammals the two hormone paralogs are AVP and OXT, with variants of AVP occurring in some species. The AVP-like hormone in birds, non-avian reptiles, amphibians and fish is AVT, and these taxa all possess an oxytocin hormone variant: mesotocin in birds, reptiles and amphibians and isotocin in fish (Stafflinger et al., 2008). These nine-amino-acid peptides are structurally very similar. AVT differs from AVP by one amino acid in the third position and from OXT by one amino acid in the eighth position (Stafflinger et al, 2008).

Mammals express three AVP receptor types (V1A, V1B and V2) and one OXT receptor which are members of the same subfamily within the rhodopsin-like family of GPCRs (Birnbaumer, 2000; Gimpl and Fahrenholz, 2001). V1A receptors have been found in all vertebrates studied, including teleost fish (Lema, 2010). In addition to mammals, V1B receptors have been identified in chickens (Cornett et al., 2003), amphibian newt species (Hasunuma et al., 2007), and the cartilaginous elephant shark (*Callorhinchus milii*; Yamaguchi et al., 2012) but have not been found in ray-finned fish. V2-type receptors involved in osmotic balance, thought previously to have arisen in tetrapods, have been identified in ray- and lobe-finned fish (Lema,

2010; Konno et al., 2009 and 2010). Subsequently, Ocampo Daza et al. (2012) and Yamaguchi et al. (2012) reported a V2BR that subdivided this lineage. Syntenic evidence supports a further split into a V2CR subtype (Lagman et al., 2013), that includes the formerly named VT1 receptor in the chicken (*Gallus gallus*) (Baeyens and Cornett, 2006). Table 1 provides a summary of the AVPR/OXTR family types, functional characteristics and taxonomic distribution.

The functional result of the hormone-receptor interaction is realized through several intermediary steps along intracellular pathways. Both the V1A and V1B receptors couple with Gq/11 protein to activate phospholipase C, inducing hydrolysis of phosphatidylinositol to inositol triphosphate (IP3) and 1,2,-diacylglycerol which increase cytoplasmic Ca^{++} levels. What had been called simply the V2 receptor in mammals, now designated V2A (see Table 1), activates the stimulatory G protein (G_S) to stimulate adenylyl cyclase increasing cAMP levels (Birnbaumer, 2000; Decaux et al., 2008). Yamaguchi et al. (2012) found that the elephant shark (*C. milii*) V2BR, although structurally similar to the teleost fish and mammalian V2AR, signals using the IP3/Ca^{++} rather than cAMP pathway. The chicken V2CR also signals through the IP3/Ca^{++} pathway (Tan et al., 2000). Therefore, the phylogenetic split among V2R subtypes is evident not only by sequence-based analyses but also by divergence of function.

Phylogenetic studies of AVP/OXT receptor sequences (Lagman et al., 2013; Yamaguchi et al., 2012; Ocampo Daza et al., 2012; Lema, 2010) have revealed a clear split between the V1R/OXTR and the V2R groups. In this study we sequenced and characterized five putative sea lamprey AVP/OXT family receptors. Our aim was to investigate whether the agnathan cyclostome sequences provide clues to the orthology within the V1R/OXTR and V2R groups, revealing the paths of duplication and divergence in the early vertebrate lineage.

2. Materials and Methods

2.1 Animals

Sea lamprey adult upstream migrants for tissue distribution studies were obtained directly from US Fish and Wildlife Service (USFWS) personnel. The USFWS captured these animals from barrier traps on the Middle River near Lake Superior in Wisconsin. Additional animals of larval, parasitic and adult life stages were obtained from USGS Hammond Bay Biological Station (HBBS) (Millersburg, MI). The larvae were maintained in sand-lined zebrafish breeder boxes within 10-gallon aquarium tanks at 15°C in a temperature-controlled room at the University of Minnesota-Duluth. A diet of moist cake yeast slurry in twice weekly feedings was provided according to the methods of Hansen et al. (1974). Parasitic-phase and adult animals were held without feeding in a 100-gallon stock tank at 15°C for a maximum of two weeks after receipt. Animals were euthanized in tricaine methanesulfonate (MS-222, Sigma, St. Louis, MO), pH 7.4, at 1g/L for larvae or 2g/L for parasitic and adult animals. Animal treatment conformed to University of Minnesota animal care standards (IACUC protocol number 1305-30612A).

2.2 Lamprey AVT receptor RNA isolation and cDNA sequence determination

Sources of RNA included larval lamprey brains, adult brains and adult gonadal tissue. Brains were pooled from several larval lampreys and immediately processed. Whole brains and gonadal tissue collected from adult male and female lampreys were stored in RNAlater (Qiagen, Germantown, MD) at -20°C prior to processing. Total RNA was isolated using TriZol® (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Homogenized tissues were passed several times through a 0.22-gauge needle prior to a QiaShredder column (Qiagen) to shear cellular components. Isolated RNA was treated with DNase I (DNA-*free*, Ambion, Austin, TX) to remove genomic DNA according to the manufacturer's instructions. The RNA

was reverse transcribed to cDNA using ThermoScript^{TM} RT (Invitrogen), with the addition of 1M betaine (Sigma) to eliminate secondary structure.

AVP/OXT receptor family sequences of teleost fish, amphibian and bird species were used in basic local alignment search tool (BLAST) queries of *P. marinus* genome database in Ensembl (current genome assembly Pmarinus_7.0). Homologous sequences identified as putative receptors were tentatively named based on original lamprey scaffold locations as Pm 807, Pm3133, Pm2017, Pm644 and Pm4769. The putative coding sequences were used to design primers for RNA isolation and cDNA sequencing (Supplementary Tables 1A-C). The Phusion High-Fidelity PCR Kit (Thermo Scientific, Waltham, MA) was used in either 50 or 25 µl reactions, with DMSO added to all reactions, HF buffer used for Pm807 and GC buffer used for all other receptors in the proportions recommended by the manufacturer. Betaine at 1M final concentration was added to all reactions except β-actin and Pm807. Final primer concentrations were 0.5 µM.

The GeneRacer^{TM} RACE kit (Invitrogen) was used in attempts to determine the UTR sequences of all putative receptor genes. Isolated larval brain RNA was dephosphorylated, decapped and ligated to the RNA oligo following the manufacturer's instructions before reverse transcribing without the addition of betaine. PCR and RACE products were visualized on 2% agarose gels with SYBR® Safe DNA Gel Stain (Invitrogen). The appropriate-sized bands were excised and purified with the QIAquick® Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. The purified products were Sanger sequenced by the University of Minnesota Genomics Center (UMGC) using an Applied Biosystems (ABI) 3730xl sequencer.

2.3 Tissue distribution patterns of AVT receptors and preprohormone

2.3.1 Parasitic Phase

Brains, pineal organ, eye, olfactory sac, gill, heart, liver, kidney, gonad, and muscle tissue were dissected from one male and one female parasitic-phase sea lamprey. All tissues were stored in individual vials of RNAlater (Ambion) at -20°C. Total RNA was isolated using TriZol® and treated with DNase I as described in section 2.2. Seven hundred nanograms of each RNA sample were reverse transcribed to cDNA using Superscript III^{TM} (Invitrogen). PCR was performed for Pm807, β-actin and AVT preprohormone in 50 µl reactions using Platinum PCR Supermix with 3 µl cDNA and a 0.5 µM final primer concentration. Primers used are listed in Supplementary Table 1B. Thermal profile for Pm807 was: 2 min at 94° followed by 34 cycles of 30 s at 94°, 30 s at 60°, and 60 s at 72°, with a 5 min elongation step at 72°. Thermal profile for β-actin was: 2 min at 94° followed by 24 cycles of 30 s at 94°, 30 s at 55°, and 60 s at 72°, with a 5 min elongation step at 72°. Thermal profile for AVT preprohormone was: 2 min at 94° followed by 35 cycles of 30 s at 94°, 30 s at 60°, and 60 s at 72°, with a 7 min elongation step at 72°. Products were visualized as in section 2.2.

2.3.2 Adult upstream migrants

Two pairs of male and female animals were collected in the migration season of separate years. Male and female pair one were collected in May 2012 and male and female pair two were collected in May 2013 from the Middle River in Wisconsin. Brains, pineal organ, eye, olfactory sac, gill, heart, liver, kidney, gonad, intestine, muscle and skin were dissected from all animals. Cloaca/penile tissues were collected from three animals of each sex obtained from HBBS in 2013 and pooled by sex. Brains collected from pair one were separated into forebrain comprised of the cerebellum, cerebrum, hypothalamus, olfactory and optic lobes; and hindbrain consisting of the medulla oblongata to approximately the start of the spinal cord (near the connection point with the spinal nerves). The brains from pair two were similarly separated into forebrain and

hindbrain, but the hypothalami were removed by gripping with forceps and slicing off the protrusion with the tip of a scalpel (Suppl. Fig. 1) and pooled by sex with those of two additional animals (see section 2.4). Pituitaries were not obtained for any of the animals. Tissue storage and RNA preparation was the same as for parasitic phase animals. For the first pair of animals: 700 ng of each RNA was reverse transcribed to cDNA using ThermoScript^{TM} RT (Invitrogen) in a 20 µl reaction; PCR was performed for Pm807, Pm3133, Pm2017, Pm644 and β-actin in a 50 µl reaction using GoTaq Master-Mix (Promega, Madison, WI), with 2 µl cDNA in the first-round reaction and 3 µl of the first-round product in the nested PCR reaction. PCR reactions for all genes except β-actin and Pm807 required a final concentration of 1M betaine.

Due to low expression in the initial PCR round for pair one, the procedure was modified for the second pair of animals as follows: 1µg of each RNA was reverse transcribed to cDNA as above but with the addition of 1M betaine (Sigma); PCR was performed in 25 µl reactions, with 2 µl cDNA in the first-round reaction and 1 µl of the first-round product in a nested PCR reaction. Thermal profiles were as above for the parasitic phase animals, with the addition of a modified thermal profile for receptor genes other than Pm807: 2 min at 95° followed by 36 cycles of 30 s at 95°, 30 s at 58°, and 60 s at 72°, with a 7 min elongation step at 72°. PCR products were visualized as described in section 2.2. Negative results were confirmed by retesting (data not shown).

2.4 Real-time quantitative PCR of forebrain and hypothalamus AVT receptor expression

RNA was isolated individually from forebrains collected from three male and three female adults, but hypothalami from these same animals were collected and pooled by sex. Primers are listed in Supplementary Table 1C. Real-time qPCR was performed on a Rotor Gene Q (Qiagen) thermocycler profile of 15 min at 95°C followed by 40 cycles of 95°C for 10 s, 60°C

for 30 s and 72°C for 60 s. QuantiTect® SYBR® Green I PCR Master Mix (Qiagen) was used in triplicate 25 µl reactions for each sample. The internal reference standard was β-actin. Assays were conducted for all genes with 0.5 µl of forebrain or hypothalamus cDNA (25 ng RNA) based on standard dilution curves conducted for Pm807 and Pm644 (Suppl. Fig. 2). Amplicon sizes ranged from 158 to 285 bp.

2.5 Sequence alignments and phylogenetic analysis

Intron-exon junctions were determined based on sequenced segments and from alignments with other species; all junctions followed the gt/ag rule (Breathnach and Chambon, 1981). Amino acid sequences were deduced from cDNA for four complete putative receptor coding sequences (Pm807, Pm3133, Pm2017, Pm644) and one partial sequence (Pm4769). These amino acid sequences were used in tblastn searches to locate the sea lamprey AVPR/OXTR gene orthologs in the Japanese lamprey (*L. japonicum*) Genome Project database, assembly version LetJap1.0 (Mehta et al., 2013). Sea lamprey and Japanese lamprey putative receptor sequences were then aligned with AVP/OXT-family receptor sequences of other species obtained from GenBank and Ensembl, using the ClustalW algorithm (Gonnet weight matrix, gap opening penalty 10.0, gap extension penalty 0.2). The alignment was used to construct a molecular phylogenetic tree using the Phylogenetic Maximum Likelihood (PhyML 3.0) method (Guindon et al., 2010) with the following settings: amino acid frequencies, proportion of invariable sites, and gamma shape parameters were estimated from the alignments, substitution rate categories was set to 8, starting tree was created using BIONJ, with NNI tree improvement, and tree topology and branch length optimization. Because we used a similar set of gene sequences to that of Lagman et al. (2013), we chose the same analysis settings, including the JTT model with 100 non-parametric bootstrap test iterations. We constructed one tree with the sea

lamprey putative AVT receptor genes and one tree without, to determine the effect of lamprey coding sequence structure on cluster confidence values. The trees were rooted using the three invertebrate octopus vasopressin-like receptors (Kanda et al., 2005) as an outgroup. The GenBank and Ensembl accession numbers for all sequences used in the analyses are listed in Suppl. Table 2.

2.6 Synteny analysis

A synteny analysis was performed to compare the sequences of neighboring genes on the lamprey scaffolds carrying AVP/OXT-family receptors to those of human (*Homo sapiens*) and chicken (*Gallus gallus*) genes. Sea lamprey (*P. marinus*) scaffolds identified in Smith et al. (2013) supplementary data and the Pmarinus_7.0 assembly in Ensembl as containing AVPR/OXTR-family genes were examined for the existence of neighboring genes. Smith and Keinath (2015) assembled near chromosome-scale linkage groups (LG) with markers anchoring the LGs to the existing scaffolds. We cross-referenced the scaffolds in the Smith et al. (2013) supplementary data with the LGs in the Smith and Keinath (2015) supplementary data.

In the searches of the Japanese lamprey database, all five of the putative receptor orthologs were found on three scaffolds: KE993674 [7.362 megabases (Mb) in length], KE993677 (6.350 Mb) and KE994228 (0.282 Mb). All genes along the entire lengths of these three scaffolds were identified by subjecting successive DNA sequence segments to NCBI Translated BLAST (blastx) searches of the human and chicken non-redundant protein sequence (nr) databases. The standard genetic code for translation and default scoring parameter algorithm settings were used. These were as follows: Matrix: BLOSUM62; Gap penalty for Existence: 11; Extension: 1; and Conditional (amino acid sequence) compositional score matrix adjustment. The output comprised possible human and chicken orthologs to each detected lamprey putative

protein-coding gene sequence. From the output, statistics including maximum bit score, E-value, percent identity and total bit score were recorded for hits with the highest total bit score among various isoforms. The bit score provides a constant statistical indicator of the alignment significance between pairs of translated gene sequence segments; the total bit score is the summed bit scores of all translated exon segments for that gene. Thus, the total bit score represents the significance of the alignment between the full-length gene in the query sequence and the subject protein, or a large proportion thereof. The lamprey DNA query lengths were adjusted to assure that the full length of a particular gene was contained within one query for an accurate total bit score. Chromosome locations of each human or chicken paralog were ascertained from the NCBI database. The FASTA amino acid sequence of high scoring human or chicken proteins were in turn used in tblastn searches of the Japanese lamprey whole genome database to find possible paralogs on other lamprey scaffolds.

3. Results

3.1 Lamprey AVT receptor gene structure and sequences

We amplified and sequenced the cDNA for the complete coding sequences of four putative AVT receptors and a partial sequence of a fifth receptor. GenBank Accession numbers are listed in Table 2. The predicted amino acid sequences are aligned in Fig. 1, where seven transmembrane (TM) domains are indicated in all of the lamprey putative AVT receptor genes. The TM domains were determined using HMMTop (Tusnády and Simon, 2001). Other characteristics typical of the rhodopsin (Class A) family of G protein-coupled receptors include the “DRY” motif in intracellular loop (ICL) 2 involved in G protein coupling (Rovati et al., 2007). Conserved residues involved in ligand binding in AVP/OXT-family receptors (Lema, 2010) are also indicated in Fig. 1. Nucleotide and deduced amino acid sequences for the receptor

genes are presented in Supplementary Figs. 3A-E. We also identified orthologs of the putative *P. marinus* receptors in the *L. japonicum* database (Mehta et al., 2013). Alignments between the amino acid sequences of the sea lamprey receptors with their orthologs in the Japanese lamprey are shown in Supplementary Figs. 4A-4E. The respective sequences of the two species are nearly identical and share the same intron/exon junction locations. Most of the differing amino acids occur in the N- and C-terminal regions.

We found a partial sequence for the AVT preprohormone on scaffold 10824 in the *P. marinus* database identical to that previously determined by Gwee et al. (2009) in the Japanese lamprey (GenBank Accession number D31871) (Suppl. Fig. 3F). We selected the β-actin gene as a positive reference gene for the expression studies and deposited a 721-bp segment in GenBank (Accession number KJ831069) (see Suppl. Fig. 3G).

A schematic diagram depicting the structure of each of the five lamprey putative AVT receptor genes and the human AVPR1A, AVPR1B, OXTR and AVPR2A genes for comparison is presented in Fig. 2. The intron locations were determined empirically in this study through cDNA sequencing, however intron lengths were not empirically determined but are based on the current genome assembly. Within the coding regions, all of the lamprey AVT receptors had an intron between the sequences putatively encoding the sixth and seventh TM domains. This was the only intron in the Pm807 and Pm4769 coding regions, while Pm2017 had one additional intron between the fifth and sixth TMs, and Pm3133 and Pm644 had an intron in that location and another intron between the second and third TM domains. RACE was used successfully to determine the sequences of the putative receptor 5' and 3' UTRs of the most highly expressed (Pm807) and the most specifically expressed (Pm2017) receptors. The 3'end of Pm4769 was also determined, however RACE attempts were unsuccessful for the 5' end of Pm4769 and for the

other receptors. The 5' UTRs in both Pm807 (Suppl. Fig 3A) and Pm2017 (Suppl. Fig. 3C) were determined to contain introns. Introns are also found in the 5' UTRs of human OXTR and AVPR1B genes (Fig. 2).

Pm807 was found to have an alternative transcription start site (TSS) 51 bp downstream of the first site, and a potential alternative translation initiation site (TIS) 24 bp downstream of the first TIS. A region of instability with a trinucleotide (CGT) repeat motif resulted in the deletion of 27 bp in the Pm2017 5' UTR of our larval lamprey brains compared with the sequence in the genome database (Suppl. Fig. 3C). No TIS could be found in the genome database for Pm4769 and attempts to find a TIS using 5' RACE were unsuccessful as noted above. Scaffold 4769 (GL481097) contains an in-frame stop codon immediately preceding the predicted open reading frame (ORF) compared to the other receptors. The Pm4769 ortholog gene found in the Japanese lamprey database has an N-terminus with a predicted translation of 26 amino acids in length from the N-terminal residue to the start of the first TM. The potential N-terminus is shown in an alignment between the deduced amino acid sequences of the Pm4769 receptor and its Japanese lamprey ortholog in Supplementary Fig. 4E, indicating a possible error in the Pmarinus_7.0 sequence construction around this gene. The last exon of Pm4769 was found on scaffold 1197 (GL477525), preventing determination of the length of the intervening intron although 3' RACE confirmed the junction between the two scaffolds. An alternative isoform of the Pm4769 mRNA was found in which a 110 bp intron is created in the first exon resulting in a frame shift (Suppl. Fig. 3E).

3.2 Tissue distribution patterns of AVT receptors and preprohormone

We determined relative tissue-, life-stage-, and sex-specific expression patterns to provide insight into each receptor's possible physiological function (Fig. 3). In our initial work in

parasitic phase animals, only Pm807, AVT preprohormone and β-actin were investigated. The cDNA bands for Pm807 were detectable from seven of eight tissues investigated (with no expression in liver) after one round of conventional reverse-transcription (RT)-PCR (Fig. 3A). Strong bands were observed for AVT in forebrain and gonads, with light or no bands detected in the other tissues. For the first pair of adult animals (2012), bands were visible for Pm807 but no bands were visible for the other putative receptors tested (Pm2017, Pm3133, and Pm644) in the initial round of PCR (data not shown). The procedure was modified for all genes in a second pair of adults collected in 2013 to include nested PCR, which also gives an accurate indication of low mRNA expression versus complete lack of expression. All receptors and the AVT preprohormone were strongly expressed in the forebrain of all animals tested (adult and parasitic) and more strongly in the hypothalamus (tested only in adults) after one round of conventional RT-PCR (Fig. 3A and Suppl. Fig. 3A).

Conventional RT-PCR products visualized on agarose gels are presented in Fig. 3A-C and Supplementary Fig. 5A and B. No bands were observed in no-template and non-reverse-transcribed negative controls in any tissue, and ribosomal bands confirmed quality of isolated RNA (Suppl. Fig. 5C). In summary, the AVT preprohormone gene was transcribed in all adult tissues that were investigated (Fig. 3B); tissues expressing all receptors after nested PCR were the gill, eye, and gonad; and among all receptors, tissues where no expression could be detected after nested PCR were more often the liver, heart and skin (Fig. 3C). Some tests with no detection in one year showed detected bands in the other, such that results between sexes differed only in a few cases (e.g., Pm644 and Pm3133 were never detected in the male heart, and Pm644 was never detected in the male skin or female liver). These differences may also have been due to individual animal condition variability.

Quantitative real-time PCR was used to assess AVT receptor mRNA levels in the hypothalamus compared with the rest of the forebrain. The hypothalamus was specifically targeted because of the known function of AVP/OXT family hormones in social and reproductive behaviors in other vertebrates. Relative receptor mRNA expression was determined in the forebrains of three males and three females individually, and hypothalami were isolated from each of these brains and pooled by sex. Forebrain reactions for Pm3133 and Pm2017 yielded C_t values >40, so relative abundances of these mRNAs were considered below the limit of reliable quantitation. Levels were similar between sexes for the other three receptors in forebrains (Fig. 4). Hypothalamus expression was substantially higher than in the rest of the forebrain for all receptors, and all receptors were detected at C_t values of 28 cycles or fewer. Pm807, Pm4769 and Pm2017 were somewhat more strongly expressed in the female than male pooled hypothalamus samples. The level of expression of Pm644 in hypothalamus was more than 20 times its expression in forebrain (Fig. 4).

3.3 Phylogenetic tree analysis

The Phylogenetic Maximum Likelihood method was employed to determine orthology relationships among the agnathan lamprey putative AVP/OXT-family receptors and those of jawed vertebrates. The amino acid sequence alignment for all receptors used to create the tree in this study is presented in Supplementary Fig. 6. The sea lamprey receptors do not align within the clades of V1A, V1B, OXT, V2A, V2B and V2C receptors established for gnathostomes. Branch support (bootstrap) values are 87% at the major branch between the V2A and V1/OXT receptor clades, 69% at the V1AR/V1BR node and 77% at the V1A/B/OXTR node. The cluster containing the sea and Japanese lamprey Pm644 and Pm4769 sequences forms a node between the V2BR and inner V2CR clusters with values less than 50%. Pm807 and Pm3133 are very

closely linked and may represent a lineage-specific duplication; they form a node with the OXTR cluster (61%) while the Pm2017 orthologs form a node with the V1BR cluster (70%). For comparison, we also constructed a phylogenetic tree without the lamprey receptors (Suppl. Fig. 7). The gnathostome receptors cluster into well-defined clades except for the V2CR genes. Bootstrap support values were higher without the lamprey sequences and were consistent with trees published by other researchers (Ocampo Daza et al., 2012, Yamaguchi et al., 2012, and Lagman et al., 2013).

3.4 Gene synteny analysis

The orthology and paralogy relationships for the AVPR/OXTR family have been previously resolved and conserved synteny demonstrated generally on four chromosome blocks in gnathostomes (Ocampo-Daza et al., 2012; Lagman et al., 2013). Our goal was to determine the neighboring genes on AVPR/OXTR-bearing LGs and/or scaffolds in both the sea and Japanese lampreys and to determine whether conserved synteny exists between these constructs in agnathans and the gnathostome syntenic chromosome blocks.

Synteny analysis results are shown in Fig. 6 and Supplementary Fig. 8. All Japanese lamprey scaffold data, full gene names and accession numbers are listed in Supplementary Table 3A (KE993674), 3B (KE993677) and 3C (KE994228). Sea lamprey scaffold 807 including the Pm807 gene was mapped (Smith and Keinath, 2015) to two different LGs (Fig. 6A, Suppl. Fig. 8). After comparison with the gene synteny we compiled for the Japanese lamprey scaffolds (Fig. 6B, Suppl. Table 3A), it is evident that scaffold 807 belongs on LG XXI (Fig. 6A) rather than LG XII (Suppl. Fig. 8). Scaffold 644 (Pm644) was also mapped to LG XXI (Smith and Keinath, 2015), similarly to the Pm807 and Pm644 orthologs in the Japanese lamprey both occurring on scaffold KE993674 (Fig. 6B). The relative orientations of sea lamprey LG XXI and Japanese

lamprey scaffold KE993674 seem to be reversed, and the order of genes jumbled although most of the same genes are present on both constructs. Sea lamprey scaffold order at particular loci on the LG is unclear as distances are given in centiMorgans (cM) rather than base pairs.

Pm3133, Pm2017, and Pm4769 are the sole protein-coding genes on their respective sea lamprey scaffolds, and neither Pm2017 nor Pm4769 could be found on any of the constructed LGs (Smith and Keinath, 2015). However, Pm2017 and Pm4769 orthologs were both found on Japanese lamprey scaffold KE993677 (Fig. 6C), which also holds interdigitated genes appearing on the human and chicken V1A and OXTR chromosomes. Only a few KE993677 genes (e.g., srgap, dnah and grip) are paralogous to genes on the KE993674 (Pm807-Pm644) scaffold. Scaffold 3133 was mapped to sea lamprey LG LVI (Smith and Keinath, 2015). The gene complement of LG LVI is depicted in Fig. 6D, showing possible paralogs to both LG XXI/scaffold KE993674 (e.g., wnt, iqsec1, foxp2/4 and lrrn) and scaffold KE993677 (e.g., nrcam/nfasc and cntn). We identified two neighboring genes near the Pm3133 gene ortholog on Japanese lamprey scaffold KE994228. These are rybp or yaf2 and pzdrn3 or 4, both genes found to be adjacent on the V1AR-V2CR chromosome block of chicken chromosome 1 (Fig. 6E). No other genes appeared to be present on the short (283 kbp) KE994228 scaffold.

In many cases only one apparent human or chicken gene exists with no paralogs in each respective genome (i.e., a 1:1 orthology). In cases with more than one possible human paralog or chicken paralog within a gene family, the paralog with the highest total bit score was counted as the most probable identity of the corresponding lamprey ortholog. In total, more than half of the genes on Japanese lamprey KE993674 mapped to the OXTR-V2BR block, on human chromosome 3 or chicken chromosome 12 (Fig. 6F). For scaffold KE993677 more of the genes mapped to the V1AR-V2CR block, on human chromosomes 12, 7 and 10 or chicken

chromosome 1 (Fig. 6F), than to any other block. However, the main caveat is that there was very little difference in total bit scores between most multiple paralog genes, making identities of these genes less certain (see Suppl. Table 3).

The tblastn paralog searches of the Japanese lamprey genome using the high-scoring human or chicken protein sequences were conducted to probe whether a second round of whole-genome duplication (2R WGD) took place in agnathans. Potential paralog hits were found to be generally scattered on many scaffolds. Forty-nine scaffolds had single hits, and 21 scaffolds had two hits (Suppl. Table 3). In addition to the longer receptor-containing scaffolds each with seven hits, other scaffolds with more than two hits are listed in Table 3. Two scaffolds (KE993818 and KE993809) had four hits and two scaffolds (KE993678 and KE993768) had five hits each.

4. Discussion

4.1 Lamprey AVT receptor gene structure and expression

We have identified and sequenced one partial and four full-length putative AVPR/OXTR family genes in the sea lamprey and utilized mRNA tissue distribution, phylogenetic maximum likelihood, and synteny analysis methods to aid in understanding AVPR/OXTR family evolution.

The intron-exon structure of the sea lamprey putative AVP/OXT receptors is shown in Fig. 2. Most notable is that the lamprey genes do not deviate from all other receptors in this family in having a conserved intron located between the sixth and seventh TM domains. This intron insertion position is conserved throughout the AVPR/OXTR family (Murasawa et al., 2009). The elephant shark AVT and OXT receptors each have only this conserved intron (Yamaguchi et al., 2012). The human AVPR2 gene has an additional mammalian-specific intron (Böselt et al., 2009) in the coding region before the first TM. Three teleost fish-specific introns in the V2AR genes are located within the sequences encoding the second, fourth and fifth TM

domains (Ocampo Daza et al., 2012). None of these additional intron positions coincide with the lamprey intron locations, which may be agnathan specific. Introns in the 5' UTRs of both Pm807 (Suppl. Fig. 3A) and Pm2017 (Suppl. Fig. 3C) parallel those insertions in the 5' UTRs of human OXTR and AVPR1B genes, but these vary in other mammals (e.g., the rat OXTR gene has only one intron insertion in the 5' UTR compared to two in humans) (Kanda et al., 2005). Thus, intron insertion positions do not appear to be reliable indicators of orthology of the different AVPR/OXTR family members among different vertebrate taxa.

Although mRNA expression does not necessarily correspond to protein expression level and function due to post-transciption regulation and protein degradation (Vogel and Marcotte, 2012), gene transcription must occur in tissues where the protein is functional. AVP/OXT family receptors have known function in the brains of jawed vertebrates. In rats, OXTR mRNA was found to be widely distributed throughout the forebrain while the V1BR which was restricted to discrete areas, especially the olfactory bulbs (Vaccari et al., 1998). The newt *Cynops pyrrhogaster* showed high expression of the V1AR in the hypothalamic nuclei and throughout the forebrain and midbrain and the V1BR showed moderate expression in only the dorsal hypothalamic nucleus and parts of the midbrain (Hasunuma et al., 2010). Our analysis detected varying sea lamprey putative receptor expression from modest in the hindbrain, high in the forebrain, to most highly expressed in the hypothalamus for all receptors (Fig. 4; Suppl. Fig. 5A and B). The AVT preprohormone was very highly expressed in all three areas (Suppl. Fig. 5A).

Of the putative V2 receptor subtypes, Pm644 mRNA was highly expressed in both males and females in the hypothalamus and much less in the forebrain (Fig. 4). As shown for both the newt (Hasunuma et al., 2010) and for the pupfish (Lema, 2010), V2A receptor expression is limited in all brain areas, while in the elephant shark hypothalamus (Yamaguchi et al., 2012),

V2BR expression was relatively high. The high Pm644 expression in the lamprey hypothalamus seems therefore more like a V2BR than a V2AR, but V2BR and V2CR expression has been studied in very few species. The other putative V2 receptor, Pm4769, displayed a similar brain-hypothalamus mRNA expression pattern.

The putative V1 receptor subtype Pm807 gene was highly expressed in all brain areas but with less difference between the forebrain and hypothalamus than for the other receptors (Fig. 4; Suppl. Fig. 5A and B). Pm807 was also consistently highly expressed in the heart, pineal organ, gills and gonads of most or all animals (Fig. 3A and B, Suppl. Fig. 5B), and in the cloaca of the 2013 adults (Fig. 3B), with moderate expression in the kidneys and no expression in liver, consistent with known functions of both V1AR and OXTR (Table 1). The Pm3133 gene was expressed in similar tissues to Pm807 but visualization required two rounds of PCR. The Pm2017 gene was expressed in the forebrain, hypothalamus, gonad, eye, gill, and weakly in the cloaca with no expression in the other tissues after nested PCR (Fig. 3B and C, Suppl. Fig. 5A and B). This limited pattern is similar to that of the isotocin receptor in the pupfish (Lema, 2010) but also reminiscent of the V1BR gene in the newt (Hasunuma et al., 2007) and the chicken (Tan et al., 2000). The possibility of conserved non-coding regulatory elements warrants further study.

4.2 AVT receptor evolutionary history

Lamprey AVT receptor genes provide new information on the potential WGD history of the AVPR/OXTR family (Figs. 5-7). Ocampo Daza et al. (2012) and Lagman et al. (2013) proposed that a vertebrate ancestral tandem gene duplication resulted in the divergence between the V1 and V2 receptor groups, and that two rounds of WGD resulted in six receptors on four chromosomes in the gnathostome ancestor, after the loss of two receptors. The 1R and 2R WGD events together could have produced eight possible receptors on four chromosomes: V1B and

lost V2; lost V1 and V2A; V1A and V2C; and OXT and V2B. However, a 1R hypothesis has been proposed based on a sea lamprey meiotic linkage map (Smith and Keinath, 2015).

Our phylogenetic analysis (Fig. 5) shows that the putative Pm807 and Pm3133 receptor amino acid sequences form a node with the OXT receptors, but are more related to each other than to the gnathostome sequences in the OXTR clade. Similarly, the Pm644 and Pm4769 sequences are more closely related to each other than to the gnathostome V2C or V2B receptors. The Pm2017 sequence forms a node between the V1AR and V1BR clades. Based only on the phylogenetic relationships we found for the lamprey putative AVP/OXT receptors, independent duplication is plausible. With the added information of conserved V1AR and OXTR neighboring genes between gnathostomes and the sea lamprey Pm807 scaffold however (Fig. 6A), a shared 1R and 2R WGD scenario becomes a more plausible alternative. Such a scenario is depicted in Figure 7A, in which the first WGD round in lampreys, or both lampreys and gnathostomes, resulted in one chromosomal block containing the V1BR and V2AR precursors, and one chromosomal block containing the V1AR/OXTR and V2BR/V2CR precursors. Employing the rationale of Smith et al. (2013), the 2R duplication of the vertebrate ancestor V1AR/OXTR-V2BR/V2CR chromosome led to paralogous formation of V1A and OXT receptor types, and that the losses of duplicate copies of neighboring genes from either chromosome differed between the lampreys and gnathostomes. The lamprey would appear to have lost the remaining V2AR gene as well as the V1BR gene, although phylogeny could indicate Pm2017 to be a V1BR. A lamprey-specific 3R duplication (Mehta et al., 2013; Nah et al., 2014) could have created the additional putative V1R subtype Pm3133 on a separate chromosome (Fig. 7A) analogous to the 3R duplication in teleost fish (Lagman et al., 2013). A lamprey-specific duplication of the GnRH-2 and -3 receptors is also proposed for the GnRH receptor family (Sower et al., 2012).

Paralogs formed on two of the four sets of 2R chromosomes were postulated to have been lost after agnathans diverged from the vertebrate lineage, creating a hidden paralogy (Kuraku, 2010).

New evidence of shared genome duplication history and orthology between cyclostomes and gnathostomes has come to light based on a recent phylogeny re-evaluation of the lamprey and hagfish in the Emx family genes (Noro et al., 2015). Previous phylogenetic analysis tended to cluster the two lamprey Emx genes with each other, supporting a cyclostome-independent duplication event. This led to the belief that neither the lamprey Emx-A nor Emx-B were orthologous to the Emx1, -2, or -3 genes in gnathostomes. Noro et al. (2015) discovered that longer homopolymeric amino acid tracts in the cyclostome gene products than in gnathostomes tended to obscure the phylogenetic outcomes. Adjusting for the homopolymer sequences and the addition of the hagfish to the analysis resulted in greater statistical support for duplication pre-dating cyclostome-gnathostome divergence. The current lack of a sequenced hagfish genome database, and other yet to be determined adjustments required to overcome cyclostome genomic idiosyncrasies (Qiu et al., 2011) still hinder lamprey gene orthology assignment on a phylogenetic basis. No AVPR/OXTR family genes have been sequenced for the hagfish, nor most of the neighboring genes on the sea lamprey scaffolds.

Our synteny analysis to determine the potential orthologies of all neighboring genes on the longer Japanese lamprey scaffolds clarifies the duplication relationship with gnathostomes. Sea lamprey Pm807 and Pm644 on linkage group (LG) XXI corresponds to Japanese lamprey scaffold KE993674. Japanese lamprey scaffold KE993677 contains both the Pm2017 and Pm4769 orthologs. Sea lamprey LG LVI contains the Pm3133 gene but no other V2R subtype. All encompass a mix of neighboring genes with seeming orthology to those on all four syntenic blocks in gnathostomes (Fig. 6). To sort relationships between agnathan and gnathostome

syntenic blocks, we designed a statistical strategy using alignment-based total bit scores, applied not just to each individual gene family but to the scaffold as a whole. Our results show that Japanese lamprey scaffold KE993674 containing the Pm807-Pm644 gene orthologs is most likely syntenic to the OXTR-V2BR block (33 of 65 human, 33 of 63 chicken genes; Fig. 6F). Scaffold KE993677 is more likely syntenic to the V1AR-V2CR block (14 of 35 human, 17 of 33 chicken genes; Fig. 6F). Therefore, orthology to gnathostome genes may not be lacking. Considering the Emx gene phylogeny results (Noro et al., 2015), the close alignment between Pm807 and Pm3133 as well as Pm644 and Pm4769 in our phylogenetic tree could be due to genomic artifacts rather than independent duplication. Homology of the LG LVI/scaffold 3133 block (Fig. 6D and E) is inconclusive because fewer genes could be analyzed on the shorter Pm3133 scaffolds. The Pm3133 gene could represent a segmental, tandem, or whole chromosome duplication that may or may not have pre-dated agnathan divergence.

Evidence of synteny with the V1AR-V2CR and OXTR/V2BR blocks supports a shared 1R scenario (Smith and Keinath, 2015) as depicted in Fig. 7B. Remnants of the ancestral V1R-V2R chromosome may have remained not only in the neighboring genes on both of the 1R chromosomes but also within the sequences of the genes themselves. After the agnathans diverged from the vertebrate lineage, a chromosomal or segmental duplication in gnathostomes may have resulted in V1A and V1B receptor subtype divergence. Phylogenetically, the Pm2017 gene may cluster with the V1BR group (Fig. 5) because of the ancestral remnants. The V1BR gene has no V2R-like partner on the same chromosome in any species yet studied. It is therefore unlikely that Pm2017 is a V1BR ortholog because it shares the scaffold with a V2B or V2C-subtype receptor, Pm4769 (Fig. 7A). Additionally, we speculate that a duplication of the V2CR gene prior to divergence of bony fishes may have evolved into the V2AR subtype. The lack of

V2CR cluster cohesiveness (Fig. 5 and Suppl. Fig. 7) could mean a V2CR re-emerged in tetrapods after the divergence of Osteichthyes.

The other 1R chromosome may have carried ancestral sequence characteristics resembling both V1AR and OXTR that have remained in the lamprey Pm807 gene (Fig. 7B). Also bearing on the trajectories of V1AR, V1BR and OXTR is the fact that tandem duplication of the AVT hormone gene to create two hormone genes (AVT and OXT) occurred after agnathan divergence and is first seen in the cartilaginous vertebrates (Gwee et al., 2009). Gwee et al. (2009) sequenced through the region that was found to contain the tandem duplication in the elephant shark, but found no evidence of a second hormone paralog in the Japanese lamprey. In our searches of both the sea and Japanese lamprey genomes we also found no evidence of an AVT hormone paralog. Therefore the new ligand specificity that may ultimately have selected for a functionally separate OXTR in gnathostomes was not present to drive fixation of nucleotide substitutions in agnathans. Based on synteny, the Pm807-Pm644 and Pm2017-Pm4769 chromosomes reflect conserved clues to a common origin that may be obscured in the phylogeny analysis by a divergent agnathan genome structural trajectory.

Although it has not yet been demonstrated that the sea lamprey putative AVT receptors are expressed as proteins and function in intracellular signaling, Uchiyama and Murakami (1994) showed AVT-dose-dependent vascular constriction leading to increased glomerular filtration rate and diuresis in Japanese lampreys. A radioimmunoassay showed that circulating AVT levels were capable of eliciting this response. The dose-response relationship was also measured for dorsal aortic blood pressure, and the responses to AVT were abolished by an AVT antagonist that worked similarly on fish and bullfrog AVT receptors (Uchiyama and Murakami, 1994). The amino acid sequences of the putative AVT receptors are nearly identical between the two

lamprey species (Suppl. Fig. 4A-E). The receptors contain residues known to be involved in ligand binding and all but Pm644 possess the “DRY” motif known to facilitate G protein coupling (Fig. 1). The tasks remain to determine which receptor is responsible for the kidney and vascular functions in these species, and in what other tissues the receptors may function.

5. Conclusions

Through mRNA expression patterning, phylogenetic evidence and gene synteny analysis we add agnathan branches to the AVP/OXT family of neuropeptide hormone receptors in vertebrate animals. Our data supports a shared 1R WGD scenario in common between lampreys and gnathostomes, with subsequent chromosomal or segmental duplications and translocations (Smith and Keinath, 2015). Due to uncertain phylogeny and the current state of genome builds in both lamprey species however, a shared 2R scenario though less plausible cannot be ruled out. Work is currently underway to demonstrate receptor function and the expression of sea lamprey receptor proteins in tissues as well as mechanisms of gene transcription regulation.

Acknowledgements

We thank Shannon RedBrook, Chris Osberg and Dr. Larry Wittmers for technical assistance, and Drs. Sara Zimmer, Kevin Diebel, Lynne Bemis, and Teresa Rose-Hellekant for reviewing the manuscript. This work is the result of research sponsored by the Minnesota Sea Grant College Program supported by the NOAA office of Sea Grant, United States Department of Commerce, under grant No. NA10OAR4170069. The U.S. Government is authorized to reproduce and distribute reprints for government purposes, notwithstanding any copyright notation that may appear hereon. This paper is journal reprint No. JR-623 of the Minnesota Sea Grant College Program.

Literature Cited

Baeyens, D.A., Cornett, L.E., 2006. The cloned avian neurohypophysial hormone receptors. Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 143, 12–19.

Birnbaumer, M., 2000. Vasopressin receptors. Trends. Endocrinol. Metab. 11, 406-410.

Böselt, I., Römpler, H., Hermsdorf, T., Thor, D., Busch, W., Schulz, A., Schöneberg, T. 2009. Involvement of the V2 vasopressin receptor in adaptation to limited water supply. PLoS ONE 4, pp. e5573.

Breathnach, R., Chambon, P., 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. 50, 349-383.

Cornett, L.E., Kirby, J.D., Vizcarra, J.A., Ellison, J.C., Thrash, J., Mayeux, P.R., et al., 2003. Molecular cloning and functional characterization of a vasotocin receptor subtype expressed in the pituitary gland of the domestic chicken (*Gallus domesticus*): avian homolog of the mammalian V1B-vasopressin receptor. Regul. Pept. 110, 231–239.

Decaux, G., Soupart, A., Vassart, G., 2008. Non-peptide arginine-vasopressin antagonists: the vaptans. Lancet 371, 1624–1632.

Ellens E.R., Kittilson J.D., Hall J.A., Sower S.A., Sheridan M.A., 2013. Evolutionary origin and divergence of the growth hormone receptor family: Insight from studies on sea lamprey. Gen Comp Endocrinol. 192, 222-236.

Gimpl, G., Fahrenholz, F., 2001. The oxytocin receptor system: structure, function, and regulation. Physiol. Rev. 81, 629-683.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst. Biol. 59, 307-321.

Gwee, P.C., Tay, B.H., Brenner, S., Venkatesh, B., 2009. Characterization of the neurohypophysial hormone gene loci in elephant shark and the Japanese lamprey: origin of the vertebrate neurohypophysial hormone genes. BMC Evol. Biol. 9, 47.

Haitina, T., Klovins, J., Takahashi, A., Löwgren, M., Ringholm, A., Enberg, J., et al., 2007. Functional characterization of two melanocortin (MC) receptors in lamprey showing orthology to the MC1 and MC4 receptor subtypes. BMC Evol. Biol. 7, 101.

Hansen, L.H., King, Jr., E.L., Howell, J.H., Smith, A.J., 1974. A culture method for sea lamprey larvae. The Progressive Fish-Culturist 36, 122-128.

Hasunuma, I., Toyoda, Kadono, Y., Yamamoto, K., F., Namiki, H., Kikuyama, S., 2010. Localization of three types of arginine vasotocin receptors in the brain and pituitary of the newt *Cynops pyrrhogaster*. Cell Tissue Res. 342, 437–457.

Hasunuma, I., Sakai, T., Nakada, T., Toyoda, F., Namiki, H., Kikuyama, S., 2007. Molecular cloning of three types of arginine vasotocin receptor in the newt, *Cynops pyrrhogaster*. Gen. Comp. Endocrinol. 151, 252–258.

Kanda, A., Satake, H., Kawada, T., Minakata, H., 2005. Novel evolutionary lineages of the invertebrate oxytocin/vasopressin superfamily peptides and their receptors in the common octopus (*Octopus vulgaris*), Biochem. J. 387, 85–91.

Konno, N., Kurosawa, M., Kaiya, H., Miyazato, M., Matsuda, K., Uchiyama, M., 2010. Molecular cloning and characterization of V2-type receptor in two ray-finned fish, gray bichir, *Polypterus senegalus* and medaka, *Oryzias latipes*. Peptides 31, 1273–1279.

Konno, N., Hyodo, S., Yamaguchi, Y., Kaiya, H., Miyazato, M., Matsuda, K., Uchiyama M., 2009. African lungfish, *Protopterus annectens*, possess an arginine vasotocin receptor homologous to the tetrapod V2-type receptor. J. Exp. Biol. 212, 2183-2193.

Kuraku, S., 2010. Paleophylogenomics of the vertebrate ancestor—Impact of hidden paralogy on hagfish and lamprey gene phylogeny. Integr. Comp. Biol. 50, 124-129.

Kuraku, S., Meyer, A., Kuratani, S., 2009. Timing of genome duplications relative to the origin of the vertebrates: Did cyclostomes diverge before or after? Mol. Biol. Evol. 26, 47-59.

Lagman, D., Ocampo Daza, D., Widmark, J., Abalo, X.M., Sundström, G., Larhammar, D., 2013. The vertebrate ancestral repertoire of visual opsins, transducin alpha subunits and oxytocin/vasopressin receptors was established by duplication of their shared genomic region in the two rounds of early vertebrate genome duplications. BMC Evol. Biol. 13, 238.

Lema, S.C., 2010. Identification of multiple vasotocin receptor cDNAs in teleost fish: Sequences, phylogenetic analysis, sites of expression, and regulation in the hypothalamus and gill in response to hyperosmotic challenge. Mol. Cell. Endocrinol. 321, 215–230.

Mehta, T.K., Ravi, V., Yamasaki, S., Lee, A.P., Lian, M.M., Tay, B.-H., et al., 2013. Evidence for at least six Hox clusters in the Japanese lamprey (*Lethenteron japonicum*). Proc. Natl. Acad. Sci. U. S. A. 110, 16044-16049.

Murasawa, S., Matsubara, H., Kijima, K., Maruyama, K., Mori, Y., Inada, M., 1995. Structure of the rat V1a vasopressin receptor gene and characterization of its promoter region and complete cDNA sequence of the 3′-end. J. Biol. Chem. 270, 20042–20050.

Nah, G.S.S., Tay, B.-H., Brenner, S., Osato, M., Venkatesh, B., 2014. Characterization of the Runx gene family in a jawless vertebrate, the Japanese lamprey (*Lethenteron japonicum*). PLoS ONE 9(11), e113445.

Noro, M., Sugahara, F., Kuraku, S., 2015. Reevaluating Emx gene phylogeny: homopolymeric amino acid tracts as a potential factor obscuring orthology signals in cyclostome genes. BMC Evol. Biol. 15, 78.

Ocampo Daza, D., Lewicka, M., Larhammar, D., 2012. The oxytocin/vasopressin receptor family has at least five members in the gnathostome lineage, including two distinct V2 subtypes, Gen. Comp. Endocrinol. 175, 135–143.

Ohno, S. (1970) Evolution by Gene Duplication (Springer, New York).

Putnam, N.H., Butts, T., Ferrier, D.E.K., Furlong, R.F., Hellsten, U., Kawashima, T., et al., 2008. The amphioxus genome and the evolution of the chordate karyotype. Nature 453, 1064-1072.

Qiu, H., Hildebrand, F., Kuraku, S., Meyer, A., 2011. Unresolved orthology and peculiar coding sequence properties of lamprey genes: the KCNA gene family as a test case. BMC Genomics 12, 325.

Rovati, G.E., Capra, V., Neubig, R., 2007. The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state. Molec. Pharmacol. 71, 959-964.

Smith, J.J., Keinath, M.C., 2015. The sea lamprey meiotic map improves resolution of ancient vertebrate genome duplications. Genome Res. 25, 1-10.

Smith, J.J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M.S., et al., 2013. Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. Nat. Genet. 45, 415-421.

Sower, S.A., Decatur, W.A., Joseph, N.T., Freamat, M., 2012. Evolution of vertebrate GnRH receptors from the perspective of a basal vertebrate. Front. Endocrinol. 3:140.

Tan, F.L., Lolait, S.J., Brownstein, M.J., Saito, N., MacLeod, V., Baeyens, D.A., Mayeux, P.R., Jones, S.M., Cornett, L.E., 2000. Molecular cloning and functional characterization of a vasotocin receptor subtype that is expressed in the shell gland and brain of the domestic chicken. Biol. Reprod. 62, 8 – 15.

Tusnády, G.E., Simon, I. 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics 17, 849-850.

Uchiyama, M., Murakami, T., 1994. Effects of AVT and vascular antagonists on kidney function and smooth muscle contraction in the river lamprey, *Lampetra japonica*. Comp. Biochem. Physiol. 107A, 493-499.

Vaccari, C., Lolait, S.J., Ostrowski, N.L., 1998. Comparative distribution of vasopressin V1B and oxytocin receptor messenger ribonucleic acids in brain. Endocrinol. 139, 5015-5033.

Vogel, C., Marcotte, E.M., 2012. Insights into the regulation of protein abundance and proteomic and transcriptomic analyses. Nat. Rev. Genet. 13, 227-232.

Yamaguchi, Y., Kaiya, H., Konno, N., Iwata, E., Miyazato, M., Uchiyama, M., et al., 2012. The fifth neurohypophysial hormone receptor is structurally related to the V2-type receptor but functionally similar to V1-type receptors. Gen. Comp. Endocrinol. 178, 519-528.

Fig. 1. Alignment and transmembrane (TM) characteristics of the sea lamprey putative AVT receptor deduced amino acid sequences. Conserved residues in all receptors are indicated by an asterisk (*), identical residues in four of five receptors, or three receptors having one residue and two another are indicated by a semicolon (:). A line above a sequence indicates each approximate TM domain. Intracellular loops (ICL) and extracellular loops (ECL) are also indicated. The "DRY" motif in ICL 2 involved in $G_{q/11}$-protein coupling is boxed. Conserved residues involved in ligand binding to the receptors are highlighted in gray.

Fig. 2. Structural schematic comparing intron-exon structure (5'-3') of the putative sea lamprey (Pm) vasotocin receptor genes to human (Hs) vasopressin and oxytocin receptor genes. Exons are indicated by boxes, grey for coding region, white for untranslated region (UTR), vertical black bars are transmembrane domains (TM), connecting lines are introns or inter-genic regions. The structures are not to scale. Sequences of UTRs with broken lines have not been determined. Complete coding sequences were determined for all putative receptors except the Pm4769, shown with grey hatching. Alternative isoforms are not shown (See Suppl. Fig. 3A-E).

Fig. 3. Tissue expression pattern of putative AVT receptor and AVT preprohormone mRNAs in (A) parasitic-phase female (F) and male (M) lampreys after one round of PCR; (B) 2013 adult female and male lampreys after one round of PCR; and (C) nested PCR of the products from the adults in panel (B). Products were visualized on 2% agarose gels with SYBR-safe® staining. Tissues examined: olfactory sac (OS), pineal organ (PO), eye, gill, heart (Hrt), liver (Liv), gonad (Gon), intestine (Int), cloaca/penis (Clo), muscle (Msc), and skin. Amplicon sizes are noted in base pairs (bp).

Fig. 4. Real-time quantitative PCR measurement of adult lamprey AVT receptor mRNA relative expression (normalized to β-actin) in forebrains (n=3 per gender, mean+SD) and hypothalami pooled by sex from the same animals. Forebrain reactions for Pm3133 and Pm2017 yielded C_t values >40.

Fig. 5. Molecular phylogenetic tree of the oxytocin and vasopressin receptor family showing sea lamprey (Pm) and Japanese lamprey (Lja) vasotocin receptors (shaded) in relation to the jawed vertebrates. The tree is rooted at the three invertebrate common octopus (Ovu) sequences as an outgroup. The tree was constructed using the Phylogenetic Maximum Likelihood (PhyML 3.0) method from a ClustalW amino acid sequence alignment. Bootstrap values based on 100 replicates are shown as percentages at the branch points (values <50% are not considered informative). The scale bar represents a genetic distance of 0.1 substitutions per site. Full scientific and common names of organisms and gene accession numbers are included in Suppl. Table 2.

Fig. 6. Analysis of neighboring genes mapped on sea lamprey (*Petromyzon marinus*) linkage groups (LG) and Japanese lamprey (*Lethenteron japonicum*) scaffolds carrying putative AVP/OXT receptor family genes. Bars in A-E with different colors represent the ratios of the

total bit scores for paralog hits on human or chicken chromosomes. Solid colored bars indicate 1:1 orthologs. Gene positions on sea lamprey LGs (A and D) are shown in centiMorgans (cM) (Smith and Keinath, 2015), with scaffold numbers and bit scores obtained from Smith et al. (2013). Scaffolds carrying AVPR/OXTR genes were located in the Japanese Lamprey Genome Project database (Mehta et al., 2013) (B, C, and E). Total bit scores for Japanese lamprey gene hits along the scaffolds were obtained in this study through blastx queries of the NCBI human and chicken non-redundant protein databases. Locations of genes are shown in megabases (Mb); orientation indicates relative direction of transcription (+ or – strand). Full gene names, accession numbers and blastx output data are given in Suppl. Table 3. (F) The paralog with the highest total bit score at each locus on Japanese lamprey scaffolds KE993674 and KE993677 was tallied to its corresponding human or chicken chromosome. See text for further explanation.

Fig. 7. Postulated whole-genome duplication (WGD) evolutionary history scenarios for the AVP/OXT receptor family and possible orthology of the lamprey AVT receptor genes. (A) 1R and 2R shared WGD scenario: The vertebrate ancestor carried a V1 and a V2 gene resulting from an invertebrate (e.g., *Ciona intestinalis*) receptor gene tandem duplication before the first WGD (1R) (Ocamo Daza et al., 2012; Lagman et al., 2013). Two duplications created four syntenic blocks in gnathostomes and the Pm807 OXTR ortholog and Pm2017 V1AR ortholog. The V1BR and other 2R duplication receptors appear to have been lost in the lamprey ("X" loci). These losses and the V1 and V2 losses in the jawed vertebrates (gnathostomes) would likely have occurred immediately after 2R, before the agnathans diverged from the vertebrate lineage. Phylogeny suggests that the Pm3133 gene is the 3R duplicate of Pm807 with all other 3R genes also lost (not shown). (B) Shared agnathan-gnathostome 1R WGD scenario: The Pm807-Pm644 genes are OXTR-V2BR orthologs, and the Pm2017-Pm4769 genes are V1AR-V2CR orthologs. A subsequent segmental duplication of the V1AR chromosome in the gnathostomes may have created the V1BR, as supported by the phylogenetic relationship between the V1AR and V1BR clades. An alternative hypothesis (not depicted) is that lampreys had one to three WGD events independently of the two rounds in gnathostomes. See text for further explanation.

Table 1. Arginine vasopressin/oxytocin receptor family members and their functional characteristics in jawed vertebrate (gnathostome) taxa.

Receptor type	**Second messenger**	**Taxa where found**[a]						**Major functions**
		M	B	R[b]	A[c]	T	C[d]	
V1A	IP3/Ca^{++}	√	√	√	√	√	√	Vascular smooth muscle vasoconstriction; social behavior, blood pressure/heartrate regulation, memory (brain); circadian rhythms (pineal); glucogenolysis (liver); uterine contraction; glomular contraction (kidney); aldosterone/cortisol secretion (adrenal)[e]
V1B (V3)	IP3/Ca^{++}	√	√	√	√		√	Adrenocorticotropic hormone (ACTH)/β-endorphin release (pituitary); stress adaptation (brain and adrenal); insulin release (pancreas)[e]
Oxytocin[f]	IP3/Ca^{++}	√	√	√	√	√	√	Uterine contraction; milk ejection (mammary); maternal behavior, male and female sexual and pro-social behavior, memory and learning (brain); seminiferous tubule contraction (testis) and sperm ejaculation (prostate)[g]
V2A	cAMP	√		√	√	√		Resorption of water in renal collecting tubules via induction and membrane insertion of aquaporin-2 (kidney); vascular smooth muscle vasodilation; sodium resorption (lung)[e]
V2B	IP3/Ca^{++}					√	√	?
V2C	IP3/Ca^{++}		√	√	√	√	√?	?

[a]**Taxa**: M-Mammals, B-Birds, R-Non-avian Reptiles, A-Amphibians, T-Teleost fish, C-Cartilaginous fish. Receptor complement based on several species except where noted. [b]Non-avian reptile receptor complement based on the anole lizard. [c]Amphibian receptor complement based on frogs and newts (see Lagman et al., 2013). [d]V1B, V2B and possible V2C subtypes in cartilaginous fish have so far been identified only in the elephant shark *Callorhinchus milii* (Yamaguchi et al., 2012); V1A and oxytocin receptors have also been detected in *Triakis scyllium* (GenBank Accession numbers BAL72722 and BAL72723, respectively). [e]Decaux et al., 2008. [f]Oxytocin receptor orthologs in birds and reptiles are called mesotocin receptors; in fish are isotocin receptors. [g]Gimpl and Fahrenholtz, 2001.

Table 2. Putative vasotocin receptor gene accession numbers and locations within the *Petromyzon marinus* genome (Assembly 7.0).

Gene/scaffold	Type	Orient.	Gene Accession	Exon	Genomic Location	Start[a]	End[b]	Contig	Contig Accession
Pm807	V1	-	KC731437	I	GL477135	241614	241741	68529	AEFG01068454
				II	GL477135	201826	202823	68527	AEFG01068452
				III	GL477135	197234	197547	68527	AEFG01068452
Pm3133	V1	-	KJ813004	I	GL479461	14318	14536	38181	AEFG01038182
				II	GL479461	11543	12009	38181	AEFG01038182
				III	GL479461	11223	11440	38182	AEFG01038183
				IV	GL479461	3826	4061	38182	AEFG01038183
Pm2017	V1	+	KJ813003	I	GL478345	15683	15709	22350	AEFG01022351
				II	GL478345	24123	24974*	22350	AEFG01022351
				III	GL478345	27413	27636	22350	AEFG01022351
				IV	GL478345	33471	33823	22350	AEFG01022351
Pm644	V2	+	KF031008	I	GL476972	181639	181986	12030	AEFG01012031
				II	GL476972	182534	183027	12030	AEFG01012031
				III	GL476972	185794	185912	12031	AEFG01012032
				IV	GL476972	188734	189032	12032	AEFG01012033
Pm4769	V2	+	KM015305	I	GL481097	22590	23534	33746	AEFG01033747
(scf. 1197)				II	GL477525	13661	14010	17058	AEFG01017059

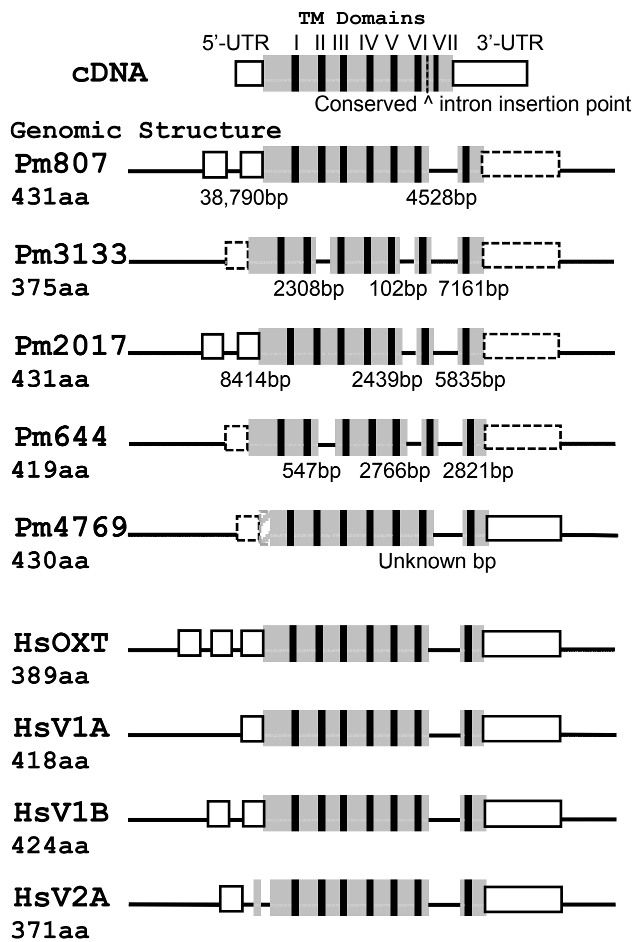
[a]Start locations are at the transcription start site for Pm807 and Pm2017, at the earliest in-frame position for Pm4769 and at the start codon for other genes. [b]End locations are at the stop codon for all genes. Orient.: relative gene orientation.

*Unstable trinucleotide repeat region resulted in the loss of 27bp. See Supplemental Figure 3A-E.

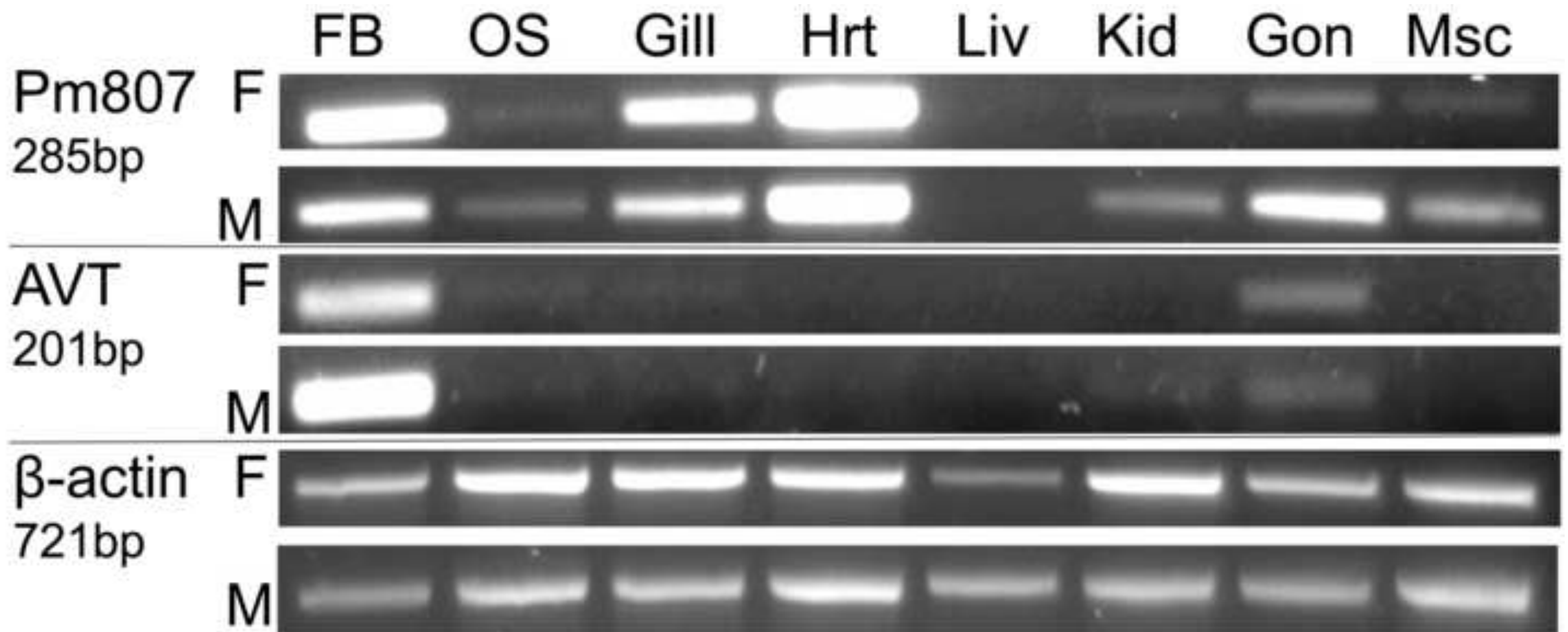
Table 3. Potential paralogs of genes neighboring the putative AVP/OXT- receptor genes occurring on other Japanese lamprey scaffolds. Scaffolds with more than two hits are shown. An additional 49 scaffolds had one hit and 21 scaffolds had two hits, either with one or two of the receptor-containing scaffolds.

Scaffold	Length (Mb)	KE993674 (Pm807/644)	KE993677 (Pm2017/4769)	KE994228 (Pm3133)
KE993674	7.362		6	1
KE993677	6.35	6		1
KE994228	0.283	1	1	
KE993818	1.533	4	--	--
KE993809	1.607	3	1	--
KE993768	2.137	2	3	--
KE994559	0.138	3	--	--
KE993837	1.305	3	--	--
KE993734	2.735	--	3	1
KE993890	0.942	2	1	--
KE994184	0.313	1	2	--
KE993813	1.565	2	2	--
KE993720	3.136	3	--	--
KE993678	6.191	1	4	--
KE993753	2.302	3	--	--

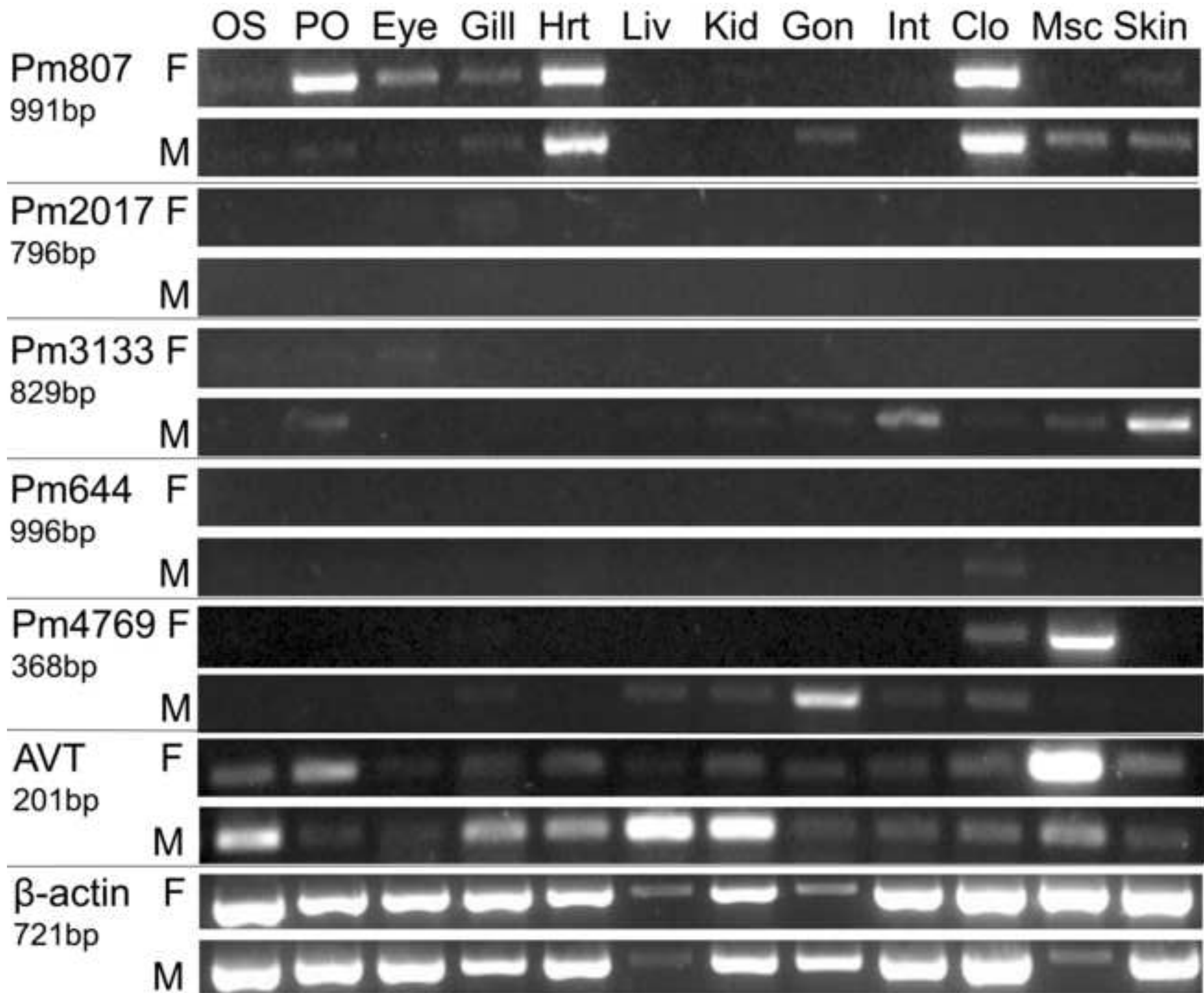
```
Pm807       MPGNVTGEMEGAILNGTGVAYAPATSNGSHVVPAEAFTAALASINATLNGSSQHALDRNE
Pm3133      ---------------------------------------MANGTANASES----TLSRDE
Pm2017      ------------------MDLAPWRNASRTPRHHNVSHAAPNAADRPEDAAAAAAAARDE
Pm644       -------------MSDDPFVGNSSDESASTLASGAWPSAAPNGTAAPHPP-LARRPERDE
Pm4769      ------------------------------------TSAAPRPGANGSAS----RNERNE
                                                  ::                *:*

                 ________TM 1___________   ICL 1   ________TM 2________
Pm807       EVAKVEIALLSIILFVAIVGNVCVLLALINTRKKTSRMHLFIMHLSIADLVVAFFQVLPQ
Pm3133      DLAKVEITILAIILCAAVVGNACVLLALLSTKKKTSRMHLFIMHLSIADLVVAFFQVLPQ
Pm2017      RLAAAEIALLAAIVAVAIAGNGSVLLALSRTRRKASRMNLFVKHLSVADLAVAAFQVLPQ
Pm644       ALARAEVTVLGLIFVLATSGNATLLVSLWRRRKHASRMHAFLVHLSVADLVVAFFQVLPQ
Pm4769      ALARAEVAVLFAMFACTTLGNSVLLTALYRRRKHASRMHVFMVHLCLADLVVAFFQVLPQ
             :* :*:: *  :   :  **  :* :*   ::::***: *: **: ***:**:******

               ECL 1        _________TM 3_________     ICL 2
Pm807       LIWKITYRFNGSDFLCRAIKYLQILGMFASTYVLIMMGIDRYIAICHPLRTLRQSS---K
Pm3133      LIWEVTFRFNGSDILCRTVKYLQILGMFASTYMLIMMAMDRYIAICHPLRTIQQSS---T
Pm2017      LSWDVTFRFRGPDALCRLVKYLQVVGMFASTYVLIAMTVDRYLAICHPLSTLRRRGEPRK
Pm644       LAWDVTDAFQASNGVCKVVKYLQVVGMFASAYMIVAMTVDRFQAVCYPMVTFKKKR---S
Pm4769      LAWDITDVFLASDFACRFIKYMQIVGMFASTYMIVGTTVDRYQAVCFPMVTFKKRM---A
            * * :*  * :::  *: :**:*::*****:*:::    **: *:* *: *

                 _________TM 4_________     ECL 2        ____________
Pm807       QSYQMIFVSWFLSMLFSLPQAFIFSMSEVEN-SGIIDCWAEFIKPWGTKAYITWMTGSVF
Pm3133      QSYMMILVSWLVSFLFSVPQAYIFSLREVQQGTGVYDCWAAFVEPWGLKAYVTWTAVAVF
Pm2017      QAHAMVACAWALSALLSTPQVVIFSLREVEE--GVFDCWADFGQPWGMRAYVTWITVSVY
Pm644       RWNGLVCAAWATSLALSTPQIFIFSLSEYEK--GIFDCTATFAAHWGAKAYVTWVTLAVF
Pm4769      YWNAPVCAVWVVSLLLSVPQAVIFAKGEVYP--GVYDCWGTFQPMWGVKAYITWMAMAIF
                :   *  * ::* **  **:  *:    *: **:: *   ** :**:** : :::

            __TM 5________       ICL 3
Pm807       IIPVIILIWCYGMITFAIWKNIKAKTQ---------------EGDSRHN-----------
Pm3133      VAPVAVLLWCYGMITAEIWRNMRAKME----------RRGGRDGRPPPKGTPARSSAA--
Pm2017      VAPVLILAACYGAI-LEICRNLRLKTR----------------GRYPKDPRGS-------
Pm644       VLPTLAVVICQAQICRIIRLNLYVKT--------------------HQGPDEEETIGGDE
Pm4769      VVPTAVLVWCQTKICRVIRLNIYFKTPQVHKDKLLLRRRKQRRRRQQKRQEDEEEVAALH
            : *:  :  *:  *   *  *:  *

                                ____________TM 6__________   ECL3
Pm807       --PAKSSAPSRVSSVRSISKAKIRTAKMTFVIIMVYIICWTPFFFVQMWSVWDSSAPFEG
Pm3133      VAPVATATASRVSSVRNISKAKIRTTKMTFVIIVVYVFCWTPFFLVQMWSVWDGSAPFEG
Pm2017      --QSQSPVCSRVSSVRCISRAKVRTIKMTLVIIVVYVVCWTPFFVIQMWAAWDETAPDDD
Pm644       RRAGGQLMPSRASSVTGISRAMINTVRMTLVIVLVYVACWAPWFTVQLWSAWDSKAPKEG
Pm4769      RGQTHAVTMTRVSSVTGVSRAMVKTVKMTLVIVIVYVLCWSPFFVAQLWSVWDPNPPFEG
                     :*:***: :*:*:: * :**:**: **: ** *:*  *:*::**  :* ::

                 __________TM 7__________
Pm807       ---IPFAIVMLLASLNSCTNPWIYMFFSGHLLYDFVRYFPCGARSRARSREADESRASDS
Pm3133      ---SAFTIVMLLASLNSCTNPWIYMFFSGHLIYDFAHCIPCCAGALRR----LETTGAGH
Pm2017      SSDPTYTIVMLLSSLNSCTNPWIYMSFSGHLLEEVMSCCRQARCATACRRQQRRRRRCRT
Pm644       ---PTFVIIMLLGNLNSCTNPWIYLWFAGGLPPGGCRGLVPHARPPAPDFNVDESVMSTA
Pm4769      ---ALFTITLLLASLNSCTNPWIYMAFSGSTPRALLSCVLCRARRGGAGGIAGGGVGSDD
                : * :** :**********: *:* :            :


Pm807       SRRNHTFVSRLTRRSLTLSSGSQHEEASSRTTSLSPVARVPKTYFA------------
Pm3133      NATDRLSATPQRHRSVTLSSCSQK----------------------------------
Pm2017      PTRRGAGSPNTEHVVVVYALSPAPTQCRSSPATAAGGTATPPETRRSAAPLWSPALP--
Pm644       TFREDRDD-----SLRGGDTPTPPFACLPPSQAQAASSLHAVSCLP------------
Pm4769      GIGGGGGTGGWVRSVPGATAHEDSMATSSVQLAGALRSQHEDDAEQTNELIRSRTISA
```



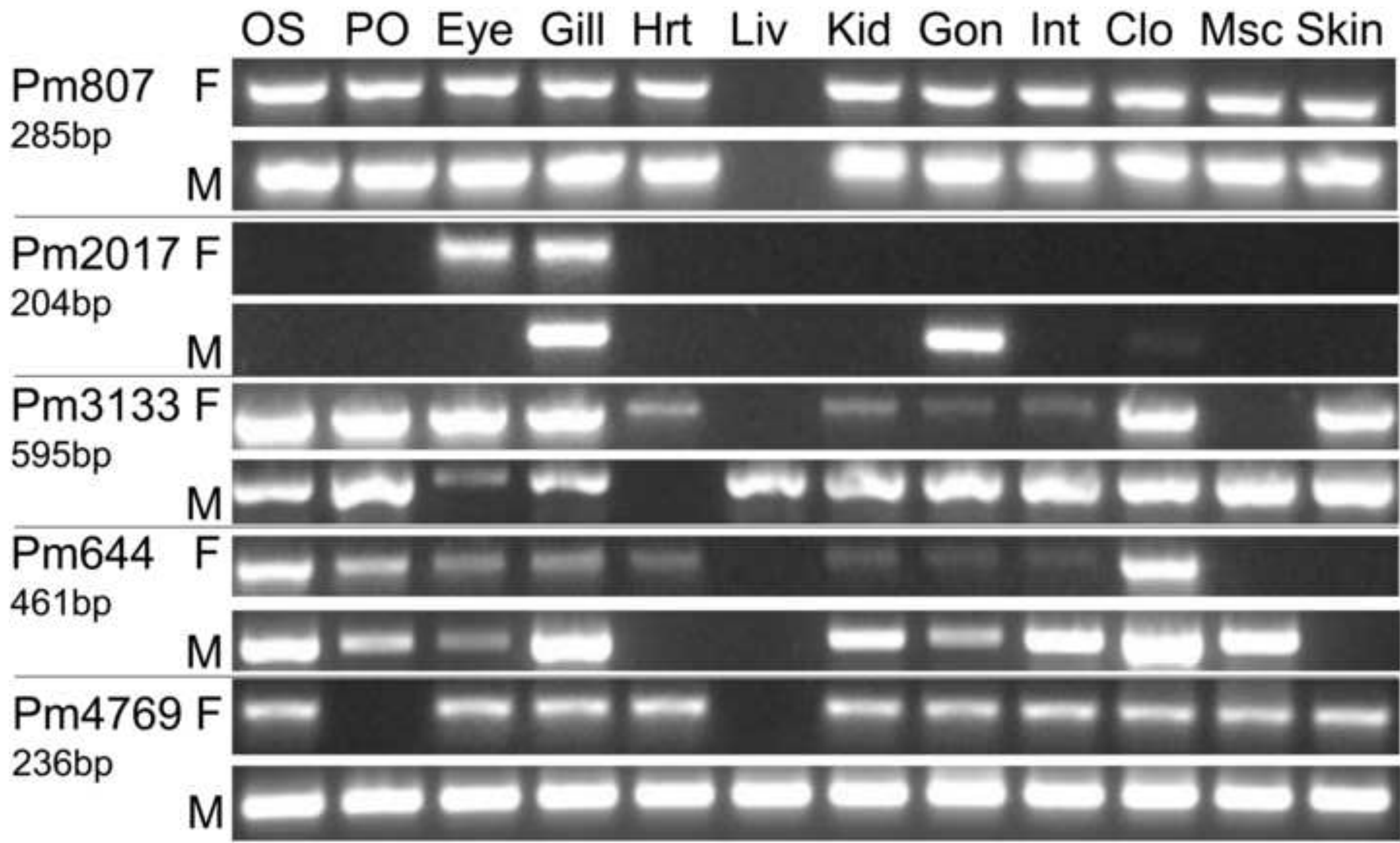
TM Domains
5'-UTR
I
II
III
IV
V
VI
VII
3'-UTR
cDNA
Conserved ^ intron insertion point
Genomic Structure
Pm807
431aa
38,790bp
4528bp
Pm3133
375aa
2308bp
102bp
7161bp
Pm2017
431aa
8414bp
2439bp
5835bp
Pm644
419aa
547bp
2766bp
2821bp
Pm4769
430aa
Unknown bp
HsOXT
389aa
HsV1A
418aa
HsV1B
424aa
HsV2A
371aa



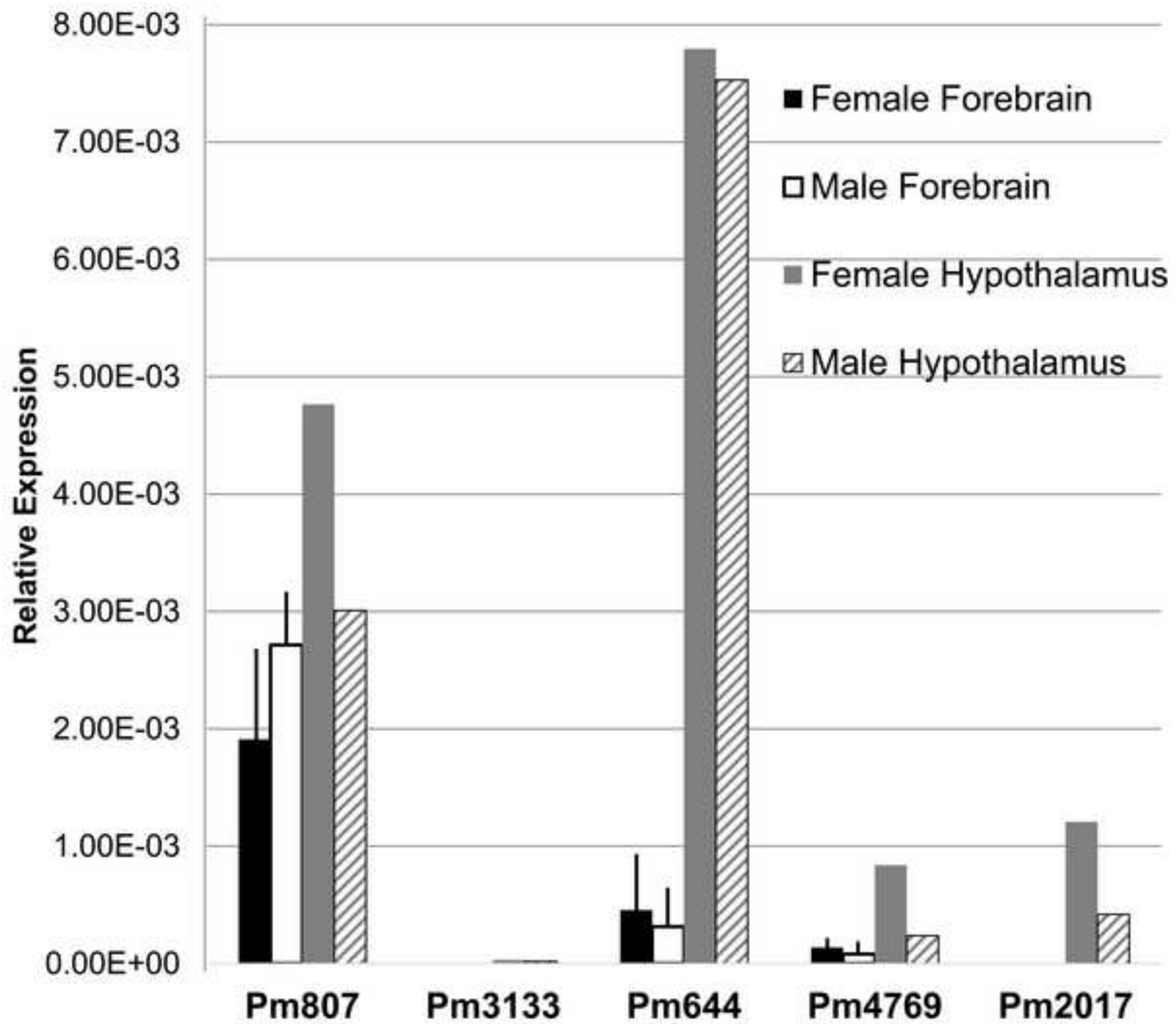
FB
OS
Gill
Hrt
Liv
Kid
Gon
Msc
Pm807
285bp
F
M
AVT
201bp
F
M
β-actin
721bp
F
M



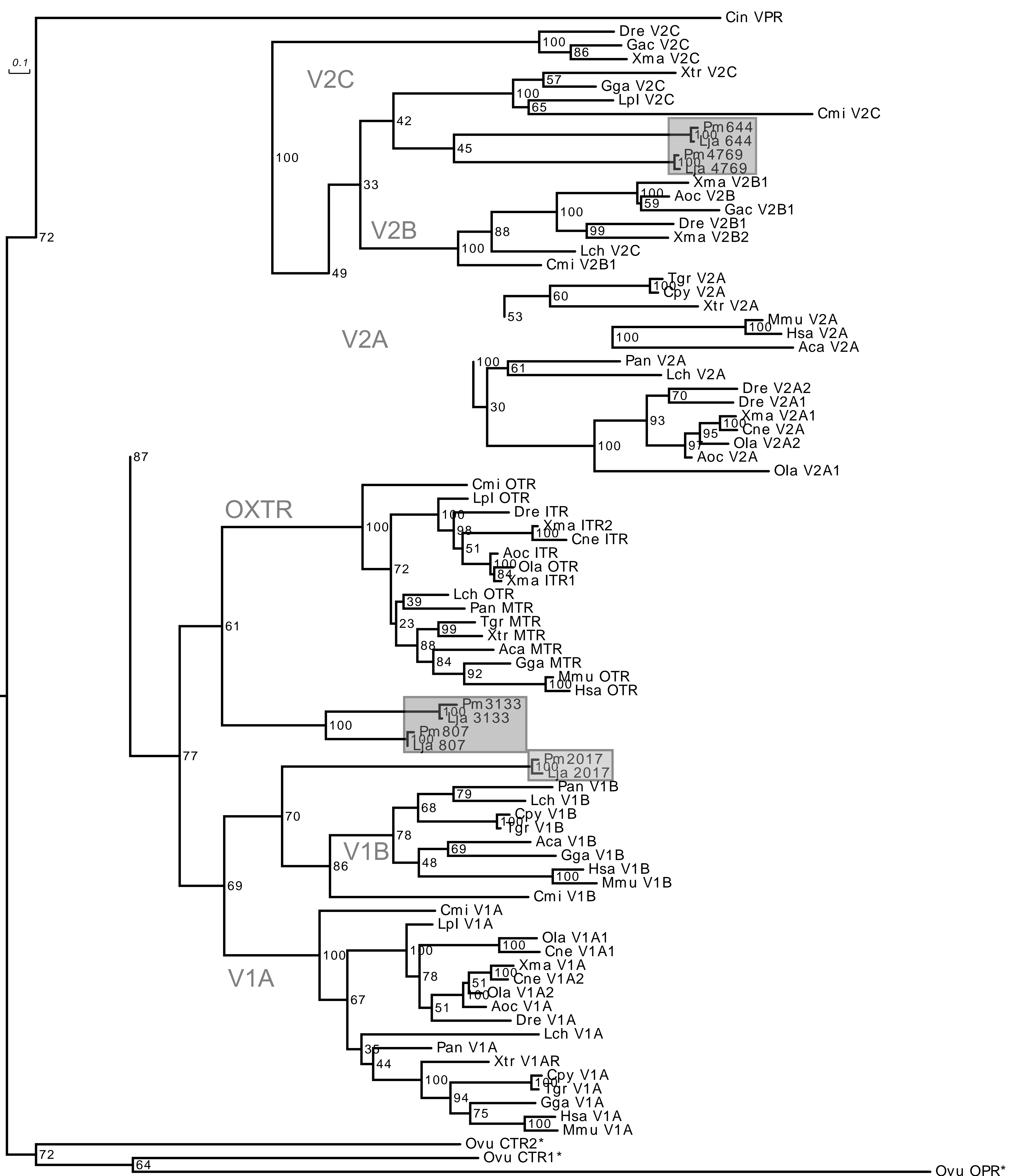
OS
PO
Eye
Gill
Hrt
Liv
Kid
Gon
Int
Clo
Msc
Skin
Pm807
991bp
Pm2017
796bp
Pm3133
829bp
Pm644
996bp
Pm4769
368bp
AVT
201bp
β-actin
721bp
F
M



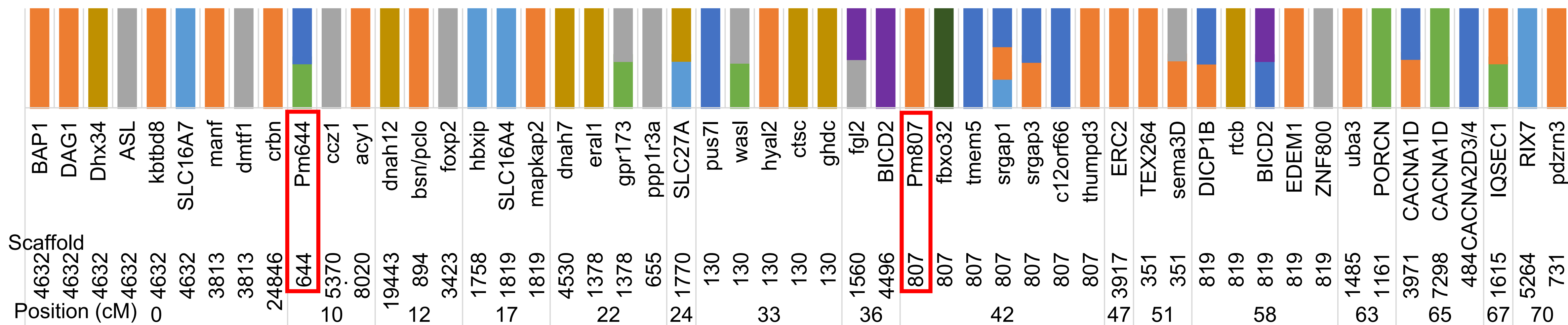
OS
PO
Eye
Gill
Hrt
Liv
Kid
Gon
Int
Clo
Msc
Skin
Pm807
285bp
F
M
Pm2017
204bp
F
M
Pm3133
595bp
F
M
Pm644
461bp
F
M
Pm4769
236bp
F
M



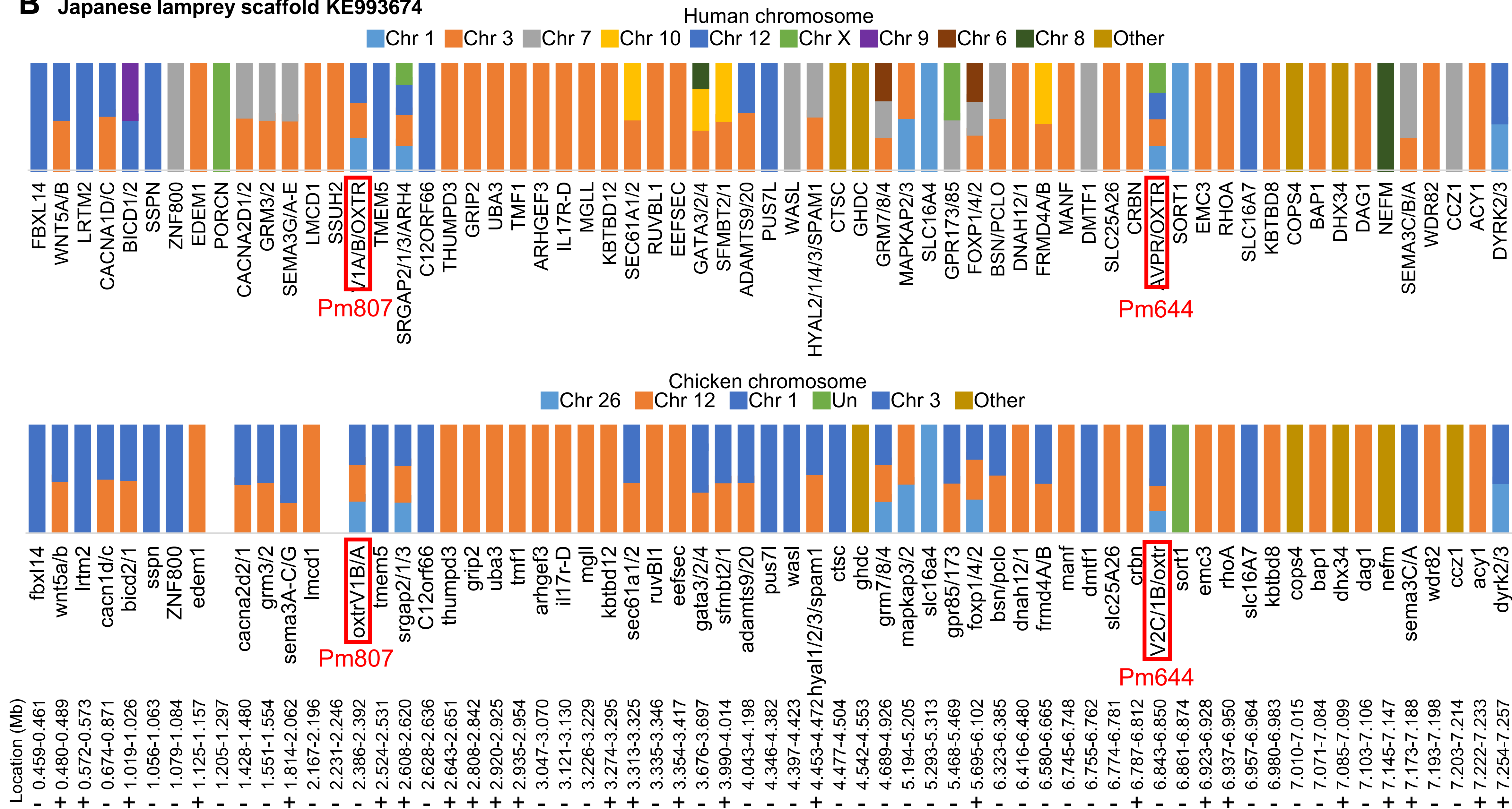
8.00E-03
7.00E-03
6.00E-03
5.00E-03
4.00E-03
3.00E-03
2.00E-03
1.00E-03
0.00E+00
Relative Expression
Female Forebrain
Male Forebrain
Female Hypothalamus
Male Hypothalamus
Pm807
Pm3133
Pm644
Pm4769
Pm2017

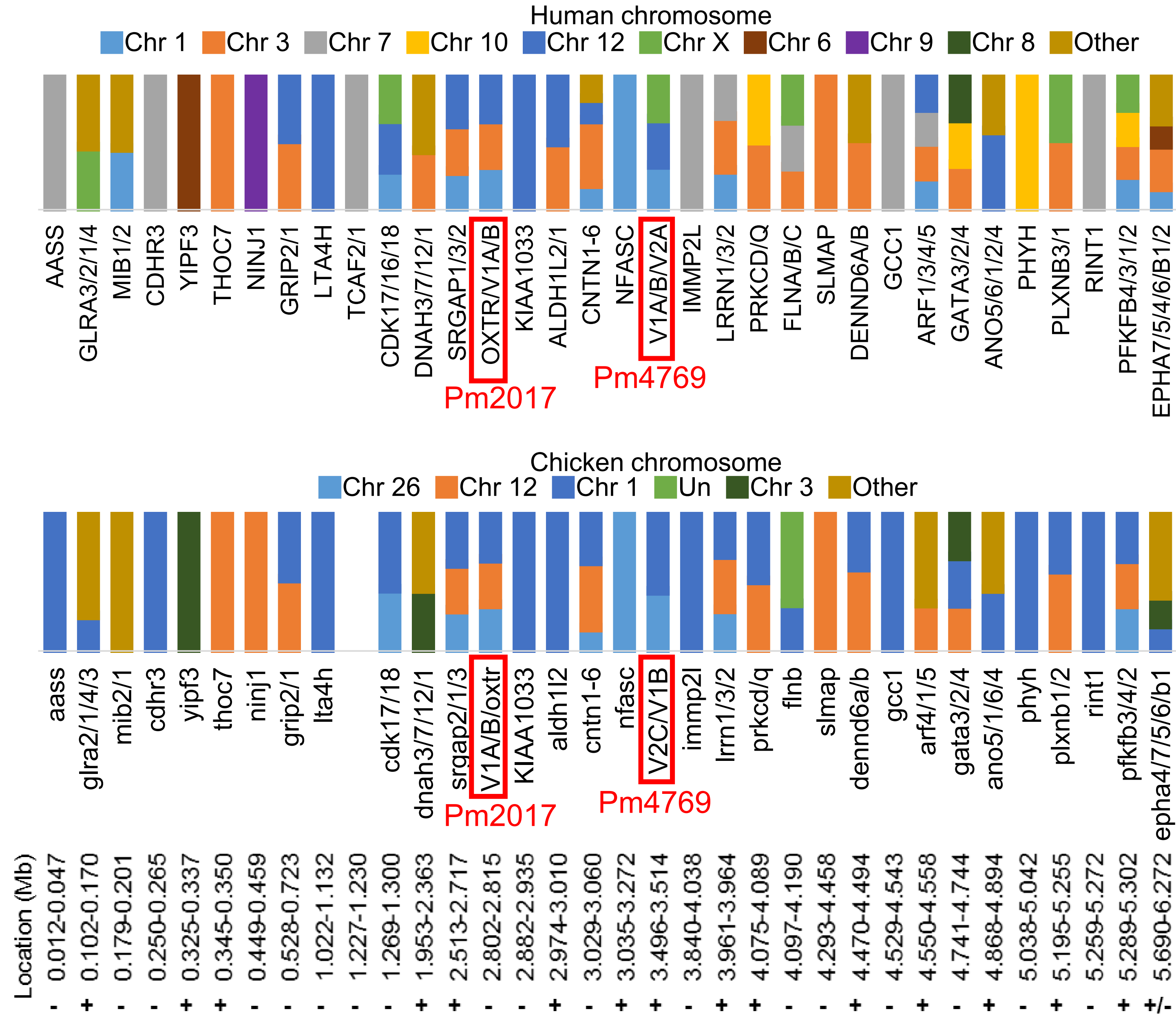


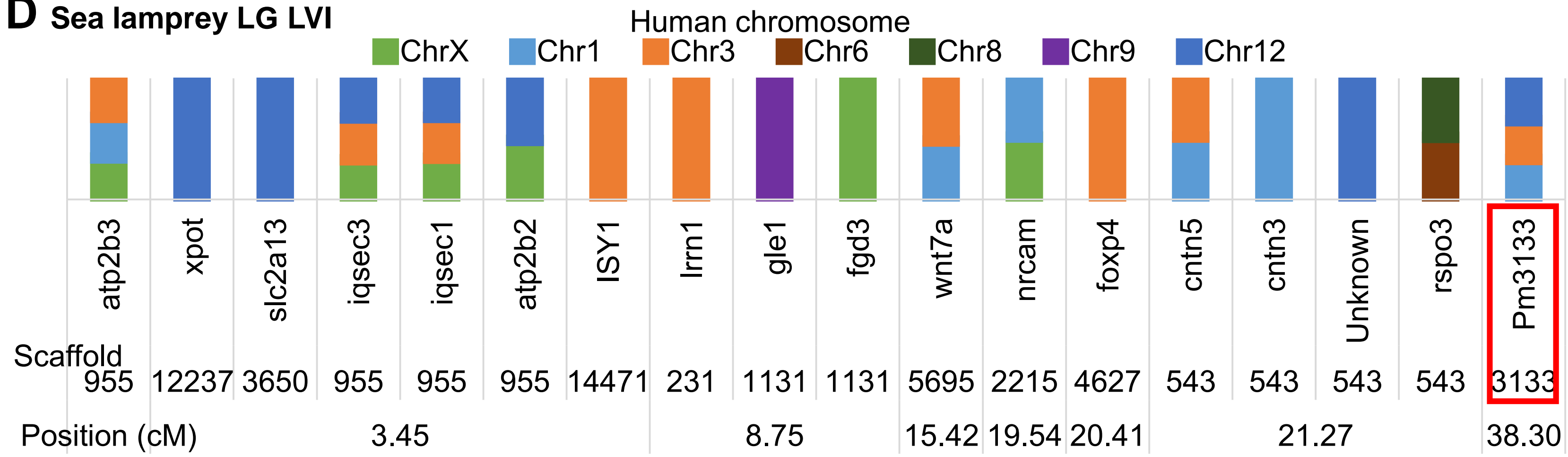
0.1
Cin VPR
Dre V2C
Gac V2C
Xma V2C
Xtr V2C
Gga V2C
Lpl V2C
Cmi V2C
Pm644
Lja 644
Pm4769
Lja 4769
Xma V2B1
Aoc V2B
Gac V2B1
Dre V2B1
Xma V2B2
Lch V2C
Cmi V2B1
V2C
V2B
V2A
Tgr V2A
Cpy V2A
Xtr V2A
Mmu V2A
Hsa V2A
Aca V2A
Pan V2A
Lch V2A
Dre V2A2
Dre V2A1
Xma V2A1
Cne V2A
Ola V2A2
Aoc V2A
Ola V2A1
OXTR
Cmi OTR
Lpl OTR
Dre ITR
Xma ITR2
Cne ITR
Aoc ITR
Ola OTR
Xma ITR1
Lch OTR
Pan MTR
Tgr MTR
Xtr MTR
Aca MTR
Gga MTR
Mmu OTR
Hsa OTR
Pm3133
Lja 3133
Pm807
Lja 807
Pm2017
Lja 2017
Pan V1B
Lch V1B
Cpy V1B
Tgr V1B
V1B
Aca V1B
Gga V1B
Hsa V1B
Mmu V1B
Cmi V1B
Cmi V1A
Lpl V1A
Ola V1A1
Cne V1A1
Xma V1A
Cne V1A2
Ola V1A2
Aoc V1A
Dre V1A
V1A
Lch V1A
Pan V1A
Xtr V1AR
Cpy V1A
Tgr V1A
Gga V1A
Hsa V1A
Mmu V1A
Ovu CTR2*
Ovu CTR1*
Ovu OPR*



A
Sea lamprey LG XXI
Human chromosome
ChrX
Chr1
Chr3
Chr7
Chr12
Chr9
Chr8
Other
BAP1
DAG1
Dhx34
ASL
kbtbd8
SLC16A7
manf
dmtf1
crbn
Pm644
ccz1
acy1
dnah12
bsn/pclo
foxp2
hbxip
SLC16A4
mapkap2
dnah7
eral1
gpr173
ppp1r3a
SLC27A
pus7l
wasl
hyal2
ctsc
ghdc
fgl2
BICD2
Pm807
fbxo32
tmem5
srgap1
srgap3
c12orf66
thumpd3
ERC2
TEX264
sema3D
DICP1B
rtcb
BICD2
EDEM1
ZNF800
uba3
PORCN
CACNA1D
CACNA1D
CACNA2D3/4
IQSEC1
RIX7
pdzrn3
Scaffold
4632
4632
4632
4632
4632
4632
3813
3813
24846
644
5370
8020
19443
894
3423
1758
1819
1819
4530
1378
1378
655
1770
130
130
130
130
130
1560
4496
807
807
807
807
807
807
807
3917
351
351
819
819
819
819
819
1485
1161
3971
7298
484
1615
5264
731
Position (cM)
0
10
12
17
22
24
33
36
42
47
51
58
63
65
67
70


B
Japanese lamprey scaffold KE993674
Human chromosome
Chr 1 Chr 3 Chr 7 Chr 10 Chr 12 Chr X Chr 9 Chr 6 Chr 8 Other
FBXL14 WNT5A/B LRTM2 CACNA1D/C BICD1/2 SSPN ZNF800 EDEM1 PORCN CACNA2D1/2 GRM3/2 SEMA3G/A-E LMCD1 SSUH2 V1A/B/OXTR TMEM5 SRGAP2/1/3/ARH4 C12ORF66 THUMPD3 GRIP2 UBA3 TMF1 ARHGEF3 IL17R-D MGLL KBTBD12 SEC61A1/2 RUVBL1 EEFSEC GATA3/2/4 SFMBT2/1 ADAMTS9/20 PUS7L WASL HYAL2/1/4/3/SPAM1 CTSC GHDC GRM7/8/4 MAPKAP2/3 SLC16A4 GPR173/85 FOXP1/4/2 BSN/PCLO DNAH12/1 FRMD4A/B MANF DMTF1 SLC25A26 CRBN AVPR/OXTR SORT1 EMC3 RHOA SLC16A7 KBTBD8 COPS4 BAP1 DHX34 DAG1 NEFM SEMA3C/B/A WDR82 CCZ1 ACY1 DYRK2/3
Pm807
Pm644
Chicken chromosome
Chr 26 Chr 12 Chr 1 Un Chr 3 Other
fbxl14 wnt5a/b lrtm2 cacn1d/c bicd2/1 sspn ZNF800 edem1 cacna2d2/1 grm3/2 sema3A-C/G lmcd1 oxtrV1B/A tmem5 srgap2/1/3 C12orf66 thumpd3 grip2 uba3 tmf1 arhgef3 il17r-D mgll kbtbd12 sec61a1/2 ruvBl1 eefsec gata3/2/4 sfmbt2/1 adamts9/20 pus7l wasl hyal1/2/3/spam1 ctsc ghdc grm7/8/4 mapkap3/2 slc16a4 gpr85/173 foxp1/4/2 bsn/pclo dnah12/1 frmd4A/B manf dmtf1 slc25A26 crbn V2C/1B/oxtr sort1 emc3 rhoA slc16A7 kbtbd8 cops4 bap1 dhx34 dag1 nefm sema3C/A wdr82 ccz1 acy1 dyrk2/3
Pm807
Pm644
Location (Mb)
- 0.459-0.461
+ 0.480-0.489
+ 0.572-0.573
- 0.674-0.871
+ 1.019-1.026
- 1.056-1.063
- 1.079-1.084
+ 1.125-1.157
- 1.205-1.297
- 1.428-1.480
- 1.551-1.554
+ 1.814-2.062
- 2.167-2.196
- 2.231-2.246
- 2.386-2.392
+ 2.524-2.531
+ 2.608-2.620
- 2.628-2.636
+ 2.643-2.651
+ 2.808-2.842
+ 2.920-2.925
+ 2.935-2.954
- 3.047-3.070
- 3.121-3.130
- 3.226-3.229
+ 3.274-3.295
+ 3.313-3.325
- 3.335-3.346
+ 3.354-3.417
- 3.676-3.697
+ 3.990-4.014
- 4.043-4.198
- 4.346-4.382
- 4.397-4.423
+ 4.453-4.472
- 4.477-4.504
- 4.542-4.553
- 4.689-4.926
- 5.194-5.205
- 5.293-5.313
- 5.468-5.469
+ 5.695-6.102
- 6.323-6.385
- 6.416-6.480
- 6.580-6.665
- 6.745-6.748
- 6.755-6.762
- 6.774-6.781
+ 6.787-6.812
- 6.843-6.850
- 6.861-6.874
+ 6.923-6.928
+ 6.937-6.950
- 6.957-6.964
- 6.980-6.983
- 7.010-7.015
- 7.071-7.084
+ 7.085-7.099
- 7.103-7.106
+ 7.145-7.147
+ 7.173-7.188
- 7.193-7.198
- 7.203-7.214
+ 7.222-7.233
+ 7.254-7.257


C
Japanese lamprey scaffold KE993677
Human chromosome
Chr 1
Chr 3
Chr 7
Chr 10
Chr 12
Chr X
Chr 6
Chr 9
Chr 8
Other
AASS
GLRA3/2/1/4
MIB1/2
CDHR3
YIPF3
THOC7
NINJ1
GRIP2/1
LTA4H
TCAF2/1
CDK17/16/18
DNAH3/7/12/1
SRGAP1/3/2
OXTR/V1A/B
KIAA1033
ALDH1L2/1
CNTN1-6
NFASC
V1A/B/V2A
IMMP2L
LRRN1/3/2
PRKCD/Q
FLNA/B/C
SLMAP
DENND6A/B
GCC1
ARF1/3/4/5
GATA3/2/4
ANO5/6/1/2/4
PHYH
PLXNB3/1
RINT1
PFKFB4/3/1/2
EPHA7/5/4/6/B1/2
Pm2017
Pm4769
Chicken chromosome
Chr 26
Chr 12
Chr 1
Un
Chr 3
Other
aass
glra2/1/4/3
mib2/1
cdhr3
yipf3
thoc7
ninj1
grip2/1
lta4h
cdk17/18
dnah3/7/12/1
srgap2/1/3
V1A/B/oxtr
KIAA1033
aldh1l2
cntn1-6
nfasc
V2C/V1B
immp2l
lrrn1/3/2
prkcd/q
flnb
slmap
dennd6a/b
gcc1
arf4/1/5
gata3/2/4
ano5/1/6/4
phyh
plxnb1/2
rint1
pfkfb3/4/2
epha4/7/5/6/b1
Pm2017
Pm4769
Location (Mb)
0.012-0.047
0.102-0.170
0.179-0.201
0.250-0.265
0.325-0.337
0.345-0.350
0.449-0.459
0.528-0.723
1.022-1.132
1.227-1.230
1.269-1.300
1.953-2.363
2.513-2.717
2.802-2.815
2.882-2.935
2.974-3.010
3.029-3.060
3.035-3.272
3.496-3.514
3.840-4.038
3.961-3.964
4.075-4.089
4.097-4.190
4.293-4.458
4.470-4.494
4.529-4.543
4.550-4.558
4.741-4.744
4.868-4.894
5.038-5.042
5.195-5.255
5.259-5.272
5.289-5.302
5.690-6.272
- + - - + + - - - - - + + - - + - + + - + + - - + - + - + - + - + +/-


D Sea lamprey LG LVI
Human chromosome
ChrX
Chr1
Chr3
Chr6
Chr8
Chr9
Chr12
atp2b3
xpot
slc2a13
iqsec3
iqsec1
atp2b2
ISY1
lrrn1
gle1
fgd3
wnt7a
nrcam
foxp4
cntn5
cntn3
Unknown
rspo3
Pm3133
Scaffold
955
12237
3650
955
955
955
14471
231
1131
1131
5695
2215
4627
543
543
543
543
3133
Position (cM)
3.45
8.75
15.42
19.54
20.41
21.27
38.30

E **Japanese lamprey scaffold KE994228**

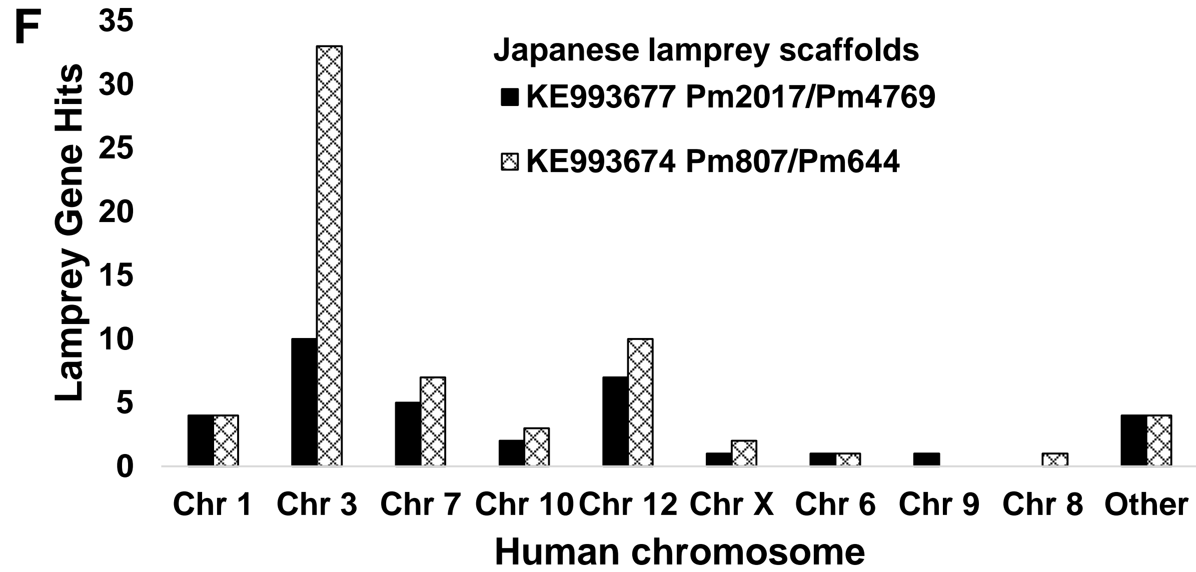
Human chromosome
Chr 1 Chr 3 Chr 12

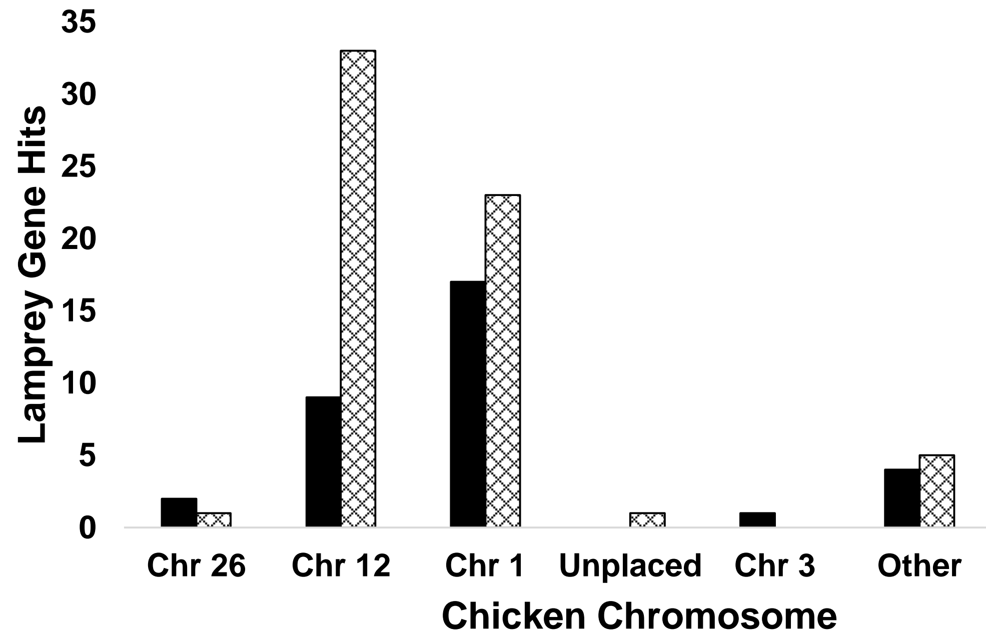
Pm3133

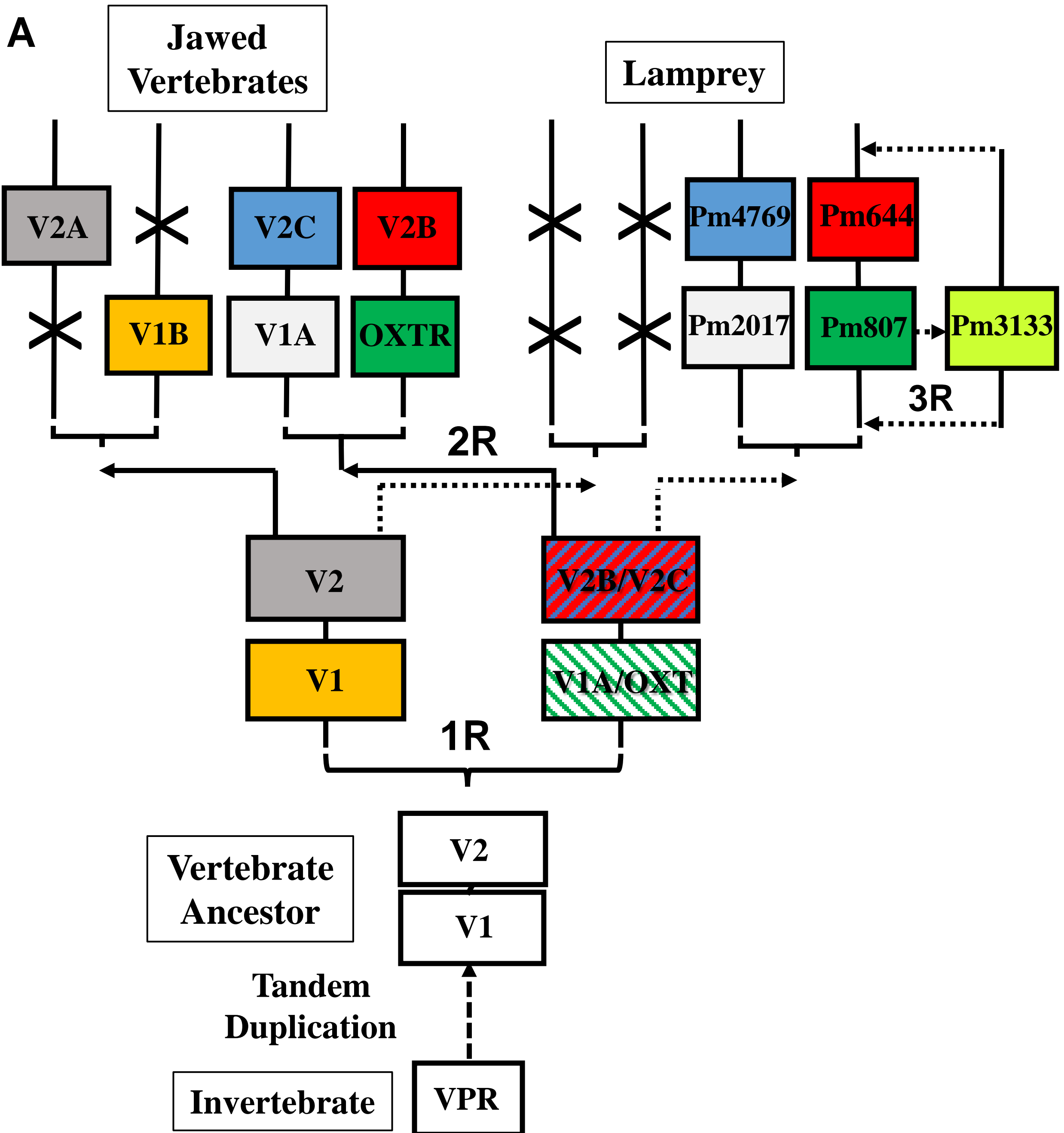
Chicken chromosome
Chr 26 Chr 12 Chr 1

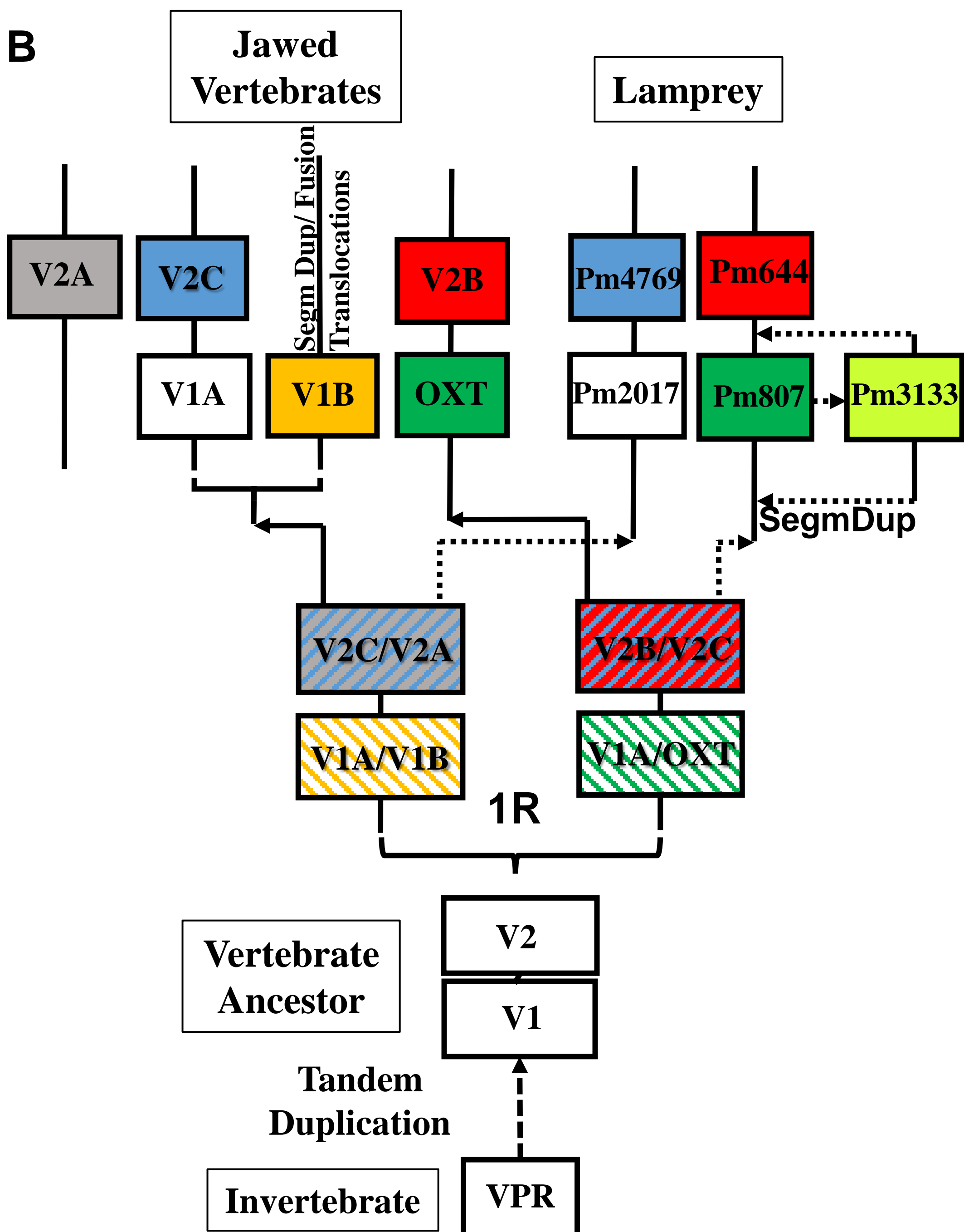
Pm3133

Location (Mb)
0.056-0.066 -
0.099-0.110 +
0.150-0.223 -


F
Lamprey Gene Hits
Japanese lamprey scaffolds
KE993677 Pm2017/Pm4769
KE993674 Pm807/Pm644
0
5
10
15
20
25
30
35
Chr 1
Chr 3
Chr 7
Chr 10
Chr 12
Chr X
Chr 6
Chr 9
Chr 8
Other
Human chromosome


Lamprey Gene Hits
0
5
10
15
20
25
30
35
Chr 26
Chr 12
Chr 1
Unplaced
Chr 3
Other
Chicken Chromosome


A
Jawed Vertebrates
Lamprey
V2A
V1B
V2C
V2B
V1A
OXTR
Pm4769
Pm644
Pm2017
Pm807
Pm3133
3R
2R
V2
V1
V2B/V2C
V1A/OXT
1R
Vertebrate Ancestor
V2
V1
Tandem Duplication
Invertebrate
VPR


B
Jawed Vertebrates
Lamprey
Segm Dup/ Fusion
Translocations
V2A
V2C
V2B
Pm4769
Pm644
V1A
V1B
OXT
Pm2017
Pm807
Pm3133
SegmDup
V2C/V2A
V2B/V2C
V1A/V1B
V1A/OXT
1R
V2
V1
Vertebrate Ancestor
Tandem Duplication
Invertebrate
VPR

Highlights: Characterization of vasotocin receptors in the sea lamprey

- We identified five putative vasotocin receptor genes in the jawless sea lamprey.
- Tissue expression patterns reveal homologies with jawed vertebrate orthologs.
- Phylogeny indicates three V1/oxytocin-clade receptors, a V2B and a V2C receptor.
- Each of two paralogous lamprey scaffolds contains a pair of syntenic V1-V2 genes.
- Synteny analysis supports a shared one-round whole-genome duplication scenario.