Studies on the origin of the hypothalamus-pituitary endocrine axis and the evolution of glycoprotein hormones in amphioxus

(ナメクジウオにおける視床下部―下垂体系の起源と
糖タンパク質ホルモンの進化に関する研究)

Thesis submitted for the degree of
Doctor of Philosophy (agriculture)
of
The University of Tokyo

December 2009

Yukiko Tando

Department of Aquatic Bioscience
Graduate School of Agricultural and Life Sciences
The University of Tokyo

Contents

Abbreviations

ACTH	adrenocorticotropic hormone
ampGPA2	amphioxus glycoprotein hormone α subunit 2
ampGPA2LP	amphioxus glycoprotein hormone α subunit 2 like protein
ampGPB5	amphioxus glycoprotein hormone β subunit 5
DDW	deionized and distilled water
DEPC	dietyl pyrocarbonate
ER	estrogen receptor
EST	expressed sequence tag
FSH	follicle-stimulating hormone
GH	growth hormone
GnRH	gonadotropin-releasing hormone
GPH	glycoprotein hormone
GTH	gonadotropin
hCG	human chorionic gonadotropin
HPA	hypothalamus-pituitary axis
IHC	immunohistochemistry
ISH	*in situ* hybridization
JGI	Joint Genome Institute
LH	luteinizing hormone
LMD	laser microdissection
NCBI	National Center for Biotechnology Information, USA
PCR	polymerase chain reaction
Pit-1	pituitary-specific transcription factor-1

POMC	proopiomelanocortin
PRL	prolactin
SL	somatolactin
SR	steroid receptor
TSH	thyroid-stimulating hormone
TRH	thyrotropin-releasing hormone
VT	vasotocin

Notes

Animal Names: Usages of animal names in this thesis followed those used in the genome database of the NCBI.

Chapter 1

General Introduction

1-1 The endocrine systems of vertebrates and invertebrates

The endocrine system is one of the major systems that control physiological processes, such as growth, metabolism, reproduction, and osmoregulation by transmission of chemical signals, or hormonal signals through their specific receptors. This regulatory system is diversified among invertebrates and vertebrates. In all vertebrates, molecular structures and physiological functions of hormones and receptors differ quite often even in the same classes, and also the presence or absence of particular endocrine organs depends on species. This diversity was mostly derived from adaptation to environments where animals survived during the history of evolution. Nonetheless, fundamental regulatory mechanisms of the endocrine systems in vertebrates can be described as almost common when compared to those in invertebrates.

In contrast to vertebrates, the endocrine systems of invertebrates are highly variable even in the same animal groups. The structure and function of their organs for endocrine regulations are much specialized, as cnidarians have scattered neurosecretory cells which secrete neuropeptides to control feeding, reproduction, and development (Mackie et al., 2003; Katsukura et al., 2004). Insects have the corpora allata and the prothoracic gland which regulate molt, metamorphosis and maturation of gonads (Moshtzky et al., 1996; Bollenbacher et al., 1975). Crustaceans have the X organ-sinus gland system and Y organ which regulate body color change and molt, and function of the androgenic gland (Keller, 1992), while octopuses have the optic gland which is involved in control of reproduction (Wells and Wells, 1969). However, most invertebrates do not have endocrine organs which are specialized for secretion of hormones. The neurosecretory systems are thus well developed in invertebrates.

1-2 The hypothalamus-pituitary axis

One of the inherent endocrine systems in vertebrates is the hypothalamus-pituitary axis

(HPA). This typical neuroendocrine axis is crucial for conversion of encephalic neural information to systemic chemical signals, which are conveyed to the target organs by bloodstream and regulate activities of target cells. This system is composed of two organs, the hypothalamus and the pituitary that are connected structurally and functionally. The hypothalamus is the basal part of the diencephalon lying below the thalamus (Swaab et al., 1993), whereas the two major subdivisions of the pituitary, adenohypophysis and neurohypophysis, locate ventrally to the brain and attached to the hypothalamus. Several hypothalamic neurosecretory centers project their axons to one of the subdivisions of the neurohypophysis, the median eminence, which is richly supplied with blood vessels that drain into the pituitary stalk, known as the hypophysial portal system (Popa and Fielding, 1930), although fishes have no median eminence and their hypothalamic neurosecretory axons project directly to the adenohypophysis.

Basically, hypothalamic neurosecretory neurons secrete small neurohormones, such as neurohypophysisal hormones, releasing hormones and inhibiting hormones. They are released into blood capillaries, travel through the portal system, and reach the pituitary where they control secretion of pituitary hormones. Pituitary hormones target certain peripheral endocrine organs such as the gonads and the thyroid which in turn release their hormones into the blood. Peripheral hormones regulate functions of target organs, and further provide either positive or negative feedback effects on the hypothalamus and the pituitary.

The two regions of the pituitary, the adenohypophysis and neurohypophysis, are derived from two different origins (Green, 1951). The adenohypophysis originates from an invagination of the oral ectoderm, or the Rathke’s pouch, beneath the diencephalon (Jacobson et al., 1979). This structure elongates to be constricted at its attachment to the oral epithelium. The adenohypophysis is thus primarily a glandular tissue, whereas the neurohypophysis originates from neuronal endings of hypothalamic neurons. The adjacent

region of the neural plate becomes a neuronal component like a funnel shaped process, which connects the Rathke's pouch that guides the neurohypophysis. The adenohypophysis can be separated into three regions: the pars distalis, the pars tuberalis, and the pars intermedia; and two subregions can be identified in the neurohypophysis, that is, the median eminence and the pars nervosa. Recently, molecular mechanisms that determine the fates of adenohypophysial endocrine cells were investigated (Scully and Rosenfeld, 2002). This study elucidated many signals that induce differentiation of cell types in the adenohypophysis.

The acquisition of the HPA is considered as a remarkable innovation in vertebrates. This endocrine system, which controls many complex endocrine functions, is consistently conserved throughout all classes of vertebrates, regardless of diverse patterns of life cycles and reproductive strategies. Such consistency in vertebrates has led an abundance of comparative researches on the mechanisms and roles of this system from the points of evolutionary views. However, why and how the HPA had emerged and acquired crucial functions only in the lineage of vertebrates are yet to be clarified.

Here, it is noteworthy that the most primitive representatives of vertebrates, hagfish and lamprey which belong to agnathan, do not have the evident median eminence. Although histochemically comparable structures of the neurohypophysis were identified in agnathan fish, neither direct neuronal innervation seen in teleosts nor vascularization seen in tetrapods is apparent between the neurohypophysis and adenohypophysis (Gorbman et al., 1983; Kobayashi and Uemura, 1972). On the basis of these classical findings, many researchers believe that a manner of hypothalamic regulation of pituitary functions is diffusion of neurohormones from the neurohypophysis to the adenohypophysis across the connective tissues (Tsukahara et al., 1986; Nozaki et al., 1994). Although it is still uncertain whether the structural characteristics of the agnathan pituitary are the origin of the HPA or not, the abovementioned findings led researches to explore homologous structures of the

hypothalamus and the pituitary in more primitive and vertebrate-related species among invertebrates such as cephalochordates and urochordates.

1-3 Amphioxus as the model for an ancestor of vertebrates

Amphioxus, which belongs to phylum Chordata, subphylum Cephalochordata, was first described by Pallas as a molluscan slug (Pallas, 1774). Cephalochordata consists of three genera (*Branchiostoma*, *Epigonichtys*, and *Asymmetron*) and widely distributed in tropical and temperate seas including Japanese coastal areas (Fig. 1-1). They are small worm-like marine animals that spend most of their lives in the sea floor, and filter-feed through the mouth (Fig. 1-2). In contrast to vertebrates, the notochord which runs along the antero-posterior axis is maintained for life. The nerve cord lies directly on the notochord. Its anterior tip is called the cerebral vesicle which shows distinctive designation (Wicht and Lacalli, 2005).

Due to its phylogenetic position as a sister group of vertebrates (Wada and Satoh, 1994), amphioxus has been considered as an important model for investigation of vertebrates. Since the finding that amphioxus has a single Hox cluster, compared to four in mammals (Garcia-Fernandez and Holland, 1994), the gene networks and developmental processes have served to illuminate the basis of morphological evolution in vertebrates through the studies of various genes (Schubert et al., 2005; Sauka-Spengler et al., 2007; Yu et al., 2007; Wada et al., 2006). The results from these researches have demonstrated that many features of the basic body plan are common between vertebrates and amphioxus. Recently, the whole-genome sequence analysis of *Branchiostoma floridae*, which live mainly in Tampa Bay, exhibited murky relationships among the three chordate groups, urochordates, cephalochordates, and vertebrates (Putnam et al., 2008), and confirmed the hypothesis proposed by Ohno (1970) that major genome duplications occurred around the origin of vertebrates.

The synteny analysis between amphioxus and human demonstrated that two rounds of

genome duplication occurred on the vertebrate stem after divergence of cephalochordates and before the split of teleosts and tetrapods (Putnam et al., 2008). Amphioxus is now recognized as the closest extant relative to the stem chordate and is the only living invertebrate that retains a vertebrate-like development and body plan throughout its lifespan (Garcia-Fernandez and Benito-Gutierrez, 2009).

1-4 Homologues of the hypothalamus and the pituitary in amphioxus

On the basis of morphological investigations, amphioxus homologues of the hypothalamus and pituitary were searched for by many researchers. Hatschek's pit, the small epithelial evagination of the oral cavity, has long been considered as the homologue of the pituitary. The most evident characteristic of this organ that supported this idea was the presence of electron-dense granules in the glandular cells of the pit (Fig. 1-3) (Tjoa and Welsch, 1974; Sahlin and Olsson, 1986). Immunoreactivity of Hatschek's pit to antisera against human gonadotropin (Chang et al., 1985; Nozaki and Gorbman, 1992) and pituitary-specific transcription factor 1 (Pit-1) (Candiani and Pestarino, 1998) showed a similarity of Hatschek's pit to the pituitary. Furthermore, gene expression of transcription factors, which are related to development of the pituitary in the preoral pit, the primodium of the Hatschek's pit (Glardon et al., 1998; Yasui et al., 2000; Boorman et al., 2002; Candiani et al., 2008), supports the idea that Hatschek's pit is homologous to the pituitary. In addition, there is a protrusion at the right side of nerve cord toward Hatschek's pit along the right side of notochord, which is considered as the homologue of the vertebrate hypothalamus (Nozaki et al., 1999). However, despite many reported studies, evidence for the presence of pituitary hormones in the pit was not substantial.

1-5 Objectives of this thesis

As previously mentioned, the HPA has crucial roles in the endocrine regulatory systems in vertebrates. Nevertheless, its origin and evolution are less well understood. I considered that researches on the evolution of the HPA will clarify primary and fundamental roles of both the hypothalamus and the pituitary, and provide substantial information on the neuroendocrine system of vertebrates including well diversified fishes. Hence, I attempted to obtain new insights into the evolution of the vertebrate HPA by use of amphioxus as a model of ancestral chordate. Circumstantial lines of evidence on morphology and development supported the similarities of the pituitary and Hatschek's pit. However, it is not yet clear whether Hatschek's pit is really the homologue of the pituitary. Similarly, it is also not known how the HPA evolved.

In this thesis, I investigated first the presence of pituitary hormone-related genes and peptides in Hatschek's pit by the sole tissue dissection method (Chapter 2). Afterward, as shown in Chapter 3, I cloned and structurally characterized glycoprotein hormone subunit genes as only pituitary hormone-related molecule found in the genome of Florida amphioxus. Chapter 4 deals with the molecular evolution of pituitary glycoprotein hormone subunits. Distribution of glycoprotein hormone subunit mRNAs in adult amphioxus is described in Chapter 5, and then distributions of gene transcripts for amphioxus homologues of hypothalamic and pituitary hormones and hormone receptors in the anterior part of amphioxus are shown in Chapter 6. Finally, I have combined altogether all the results and taken further consideration for the evolution of the HPA in Chapter 7.

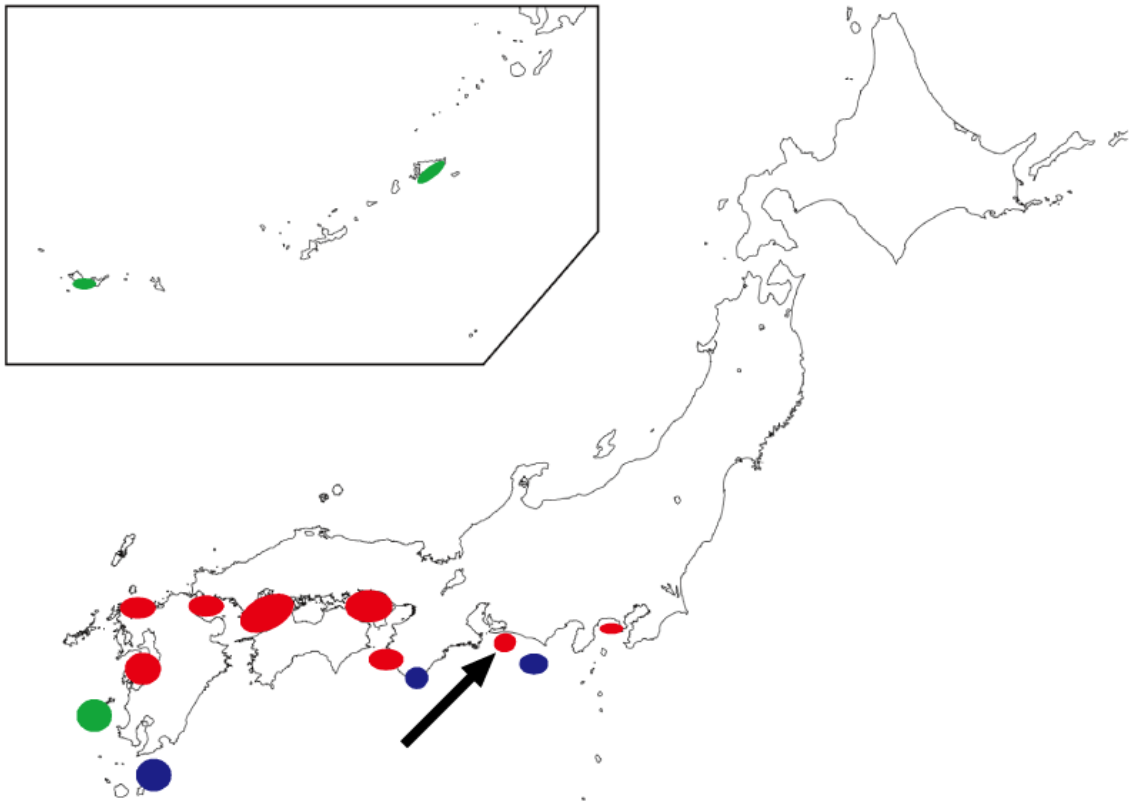

Figure 1-1. Major habitats of lancelets (Cephalochordata) in Japan. The inset shows the Okinawa islands in southern Japan. Cephalochordata is composed of three genera, *Branchiostoma* (habitats of which are indicated by red spots), *Epigonichthys* (blue spots), and *Asymmetron* (green spots). Experimental animals of *Branchiostoma belcheri* were collected at the coastal area off Atsumi Peninsula (arrow).

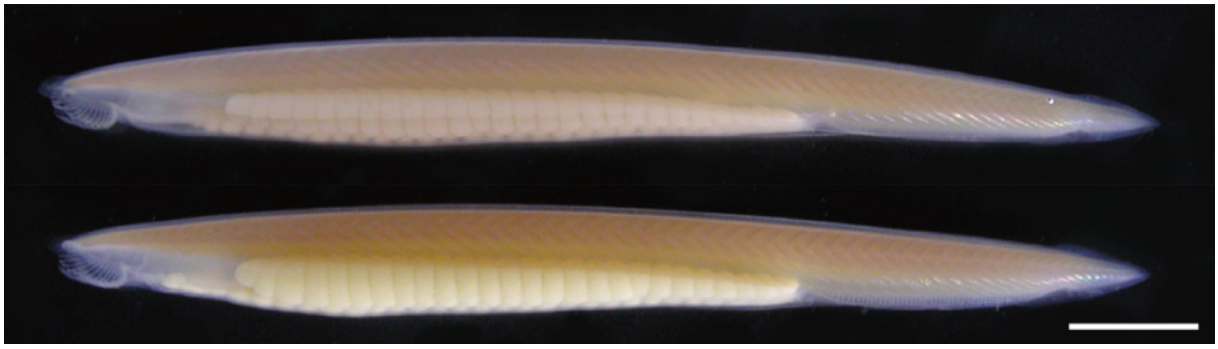

Figure 1-2. Photograph of male and female *Branchiostoma belcheri*. In breeding season, males have whitish testes (upper photograph), while females have yellowish ovaries (lower photograph). Other regions are similar between both sexes. Scale bar, 5 mm.

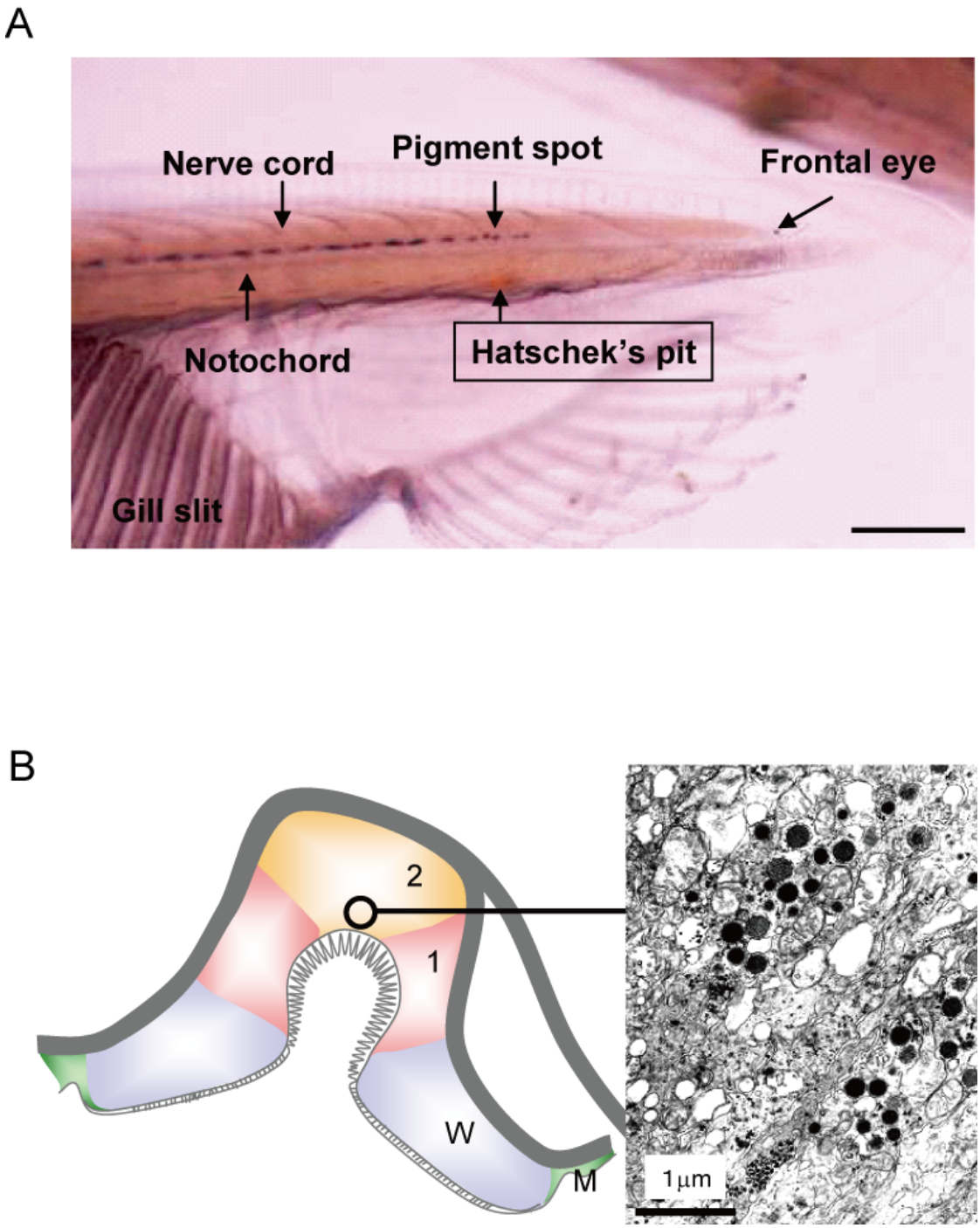

Figure 1-3. Structure of Hatschek's pit. (A) Light microscopic view of the right side of head region of amphioxus. Hatschek's pit is located at the dorsal side of oral cavity. Scale bar, 1mm. (B) Transmission electron micrographs of a region. The right panel shows the presence of electron dense granules in the apical part of type 2 groove cell (Terakado, unpublished data). Types of Hatschek's pit cells are shown with different colors (schematic drawing of Sahlin and Olsson (1986) was modified). 1, type 1 groove cells; 2, type 2 groove cells; M, margin cells; W, wheel organ cells.

Chapter 2

Survey of Homologous Genes for Vertebrate Adenohypophysial Hormones from Amphioxus

2-1 Introduction

Hatschek's pit is a groove-like structure in the mouth cavity, and locates between myomeres 3 and 4 (Fig. 2-1). This evagination extends dorsad, but does not reach the nerve cord. On the basis of its structure and developmental morphology, Hatschek's pit has long been considered as the homologue of the vertebrate pituitary. The discovery of electron dense granules in glandular cells (Tjoa et al., 1974), as well as immunopositive reactions to gonadotropin (GTH) and gonadotropin-releasing hormone (GnRH) in Hatschek's pit (Chang et al., 1982), brought attention to the function of Hatschek's pit. Upon a detailed histological analysis, Sahlin and Olsson (1986) observed many granules and developed apical mitochondria in the pit, and suggested its endocrine function. Following the finding of Chang et al. (1982), Nozaki and Gorbman (1992) confirmed weak immunopositive staining in Hatschek's pit with antisera against human chorionic gonadotropin β (hCGβ) and luteinizing hormone β (LHβ). Immunoreactivity to pituitary specific transcription factor (Pit-1) was also found in the pit (Candiani and Pestarino, 1998).

The similarity of developmental morphology between Hatschek's pit and the vertebrate pituitary raised the concept of homology between these two organs. Hatschek's pit is formed by dorsal evagination of mouth cavity (Goodrich, 1918). This developmental process and its morphology are similar to the proportion of Rathke's pouch and the invagination of the placode of the anterior pituitary in the oral ectoderm. The placode extends toward the ventral diencephalon in early developmental stages of the vertebrate pituitary (Schwind, 1928). The expression patterns of genes related to pituitary development in amphioxus larvae further support the idea of similarity between Hatschek's pit and pituitary. Actually, expression of *Pax6*, (Glardon et al., 1998), *Pitx* family genes (Yasui et al., 2000; Boorman and Shimeld, 2002), and *Pit-1* (Candiani et al., 2008) were detected in the preoral pit, the primordium of the Hatschek's pit. Besides these findings, Nozaki et al (1999) and Gorbman

et al (1999) noted that the nerve cord extends a protrusion to the Hatschek's pit along the right side of notochord, and suggested that the nerve protrusion and Hatschek's pit are homologous to the hypothalamo-adenohypophysial structure in vertebrates.

Extracts of Hatschek's pits stimulated gonadal function in amphioxus, and further in toads (Fang and Wang, 1984), suggesting that Hatschek's pit contains biologically active vertebrate-like GTH. The presences of vertebrate-like sex steroidogenesis (Mizuta and Kubokawa, 2007) and four types of GnRH receptor (Tello and Sherwood, 2009) further provided an implication of the presence of the hypothalamus-pituitary-gonad axis seen in all classes of vertebrates. However, until now, substantial evidence is not provided for the presence of GTH and other pituitary hormones in Hatschek's pit nor in any portion in amphioxus.

As mentioned above, the majority of investigation for the function of Hatschek's pit is based on histological studies, except for the *in vivo* study by Fang and Wang (1984). The most probable reason for the lack of physiological studies is that isolation of Hatschek's pit is quite difficult. The size of Hatschek's pit is only 100 μm in length and 50 μm in width. In addition, the half of their ridge is tightly associated with connective tissue of the edge of muscles. Therefore, when we isolate Hatschek's pit with forceps under a microscope, its structure will be broken by manipulation and muscle tissues are likely to be contaminated. Such difficulty prevents direct approach to investigate the pit by extraction and isolation of hormones or gene transcripts from them.

Recently, a laser microdissection (LMD) technique has been well established to isolate individual cells or subcellular structures from a heterogeneous cell population such as neurons in the brain and gamete cells in the gonad tissues (Ladanyi et al., 2006). The LMD has been successfully coupled with molecular biological techniques by which researchers analyze the genome, transcriptome, and proteome of microdissected cells and tissues.

In the experiment reported in this chapter, I aimed to survey pituitary hormone homologues in Hatschek's pit. For this investigation, I used a LMD technique for isolation of Hatschek's pit. Then, I analyzed gene expression profiles of Hatschek's pit by combining a subtractive hybridization and an expressed sequence tag (EST) analysis of cDNA library from microdissected Hatschek's pits. In addition, analysis of micro-proteomics of Hatschek's pits was carried out to confirm a presence of peptide hormones. Taken together, the function of Hatschek's pit was discussed from the present results.

2-2 Materials and Methods

2-2-1 Animals

Mature adult amphioxus, *B. belcheri*, were collected in the Enshu Nada Sea, Japan, during the breeding season (Kubokawa et al., 1998; Kubokawa et al., 2003), and taken to the Ocean Research Institute, the University of Tokyo. They were maintained in seawater tanks at about 25 ^{o}C until use (Kubokawa et al., 2003; Mizuta and Kubokawa, 2004).

2-2-2 Dissection of Hatschek's pit by LMD

The head region of amphioxus was removed, rapidly frozen at -80^{o}C, and embedded in OCT compound (Ted Pella, Redding, CA). Frozen heads were transversally cut at 14 μm with a cryostat (Leica, Wetzlar, Germany), and dried on an RNase-free glass slide with film (Matsunami Glass, Osaka, Japan). About 50 sections, which included Hatschek's pit, on a single glass slide were treated with 75% ethanol in deionized and distilled water (DDW) containing 0.1% dietyl pyrocarbonate (DEPC) on ice for 3 min. After a wash with DEPC DDW for 10 sec, the sections were stained with 1% cresyl violet for 1 min, and washed also with DEPC DDW for 2 min. Following the wash, the sections were dehydrated serially with

95% and 100% ethanol each for 30 sec, and then two times of xylene each for 2 min. The region of Hatschek's pit was dissected out with an AS Laser Microdissection Microsytem (Leica) (Figs. 2-2, 2-3) and collected into a lid of 0.5 ml tube which was filled with 30 μl of ISOGEN (Nippon gene, Tokyo, Japan). The solution which contained a piece of tissue was collected and kept at -80^{o}C until RNA was extracted.

2-2-3 Subtractive hybridization of Hatschek's pit dissected by LMD

A total of 140 sections that included Hatschek's pit were removed by LMD. Muscle tissues for subtractive hybridization were similarly dissected out under a microscope, frozen by liquid nitrogen, and stored at -80^{o}C until use. Frozen muscle tissues were homogenized with Lysing Matrix D (Qbiogene, Irvine, CA) and Fastprep FP100 (Qbiogene), and used for extraction of total RNA. Total RNAs from Hatschek's pits and muscles were extracted with ISOGEN according to the manufacture's instruction, and dissolved in 20 μl of DEPC DDW. These total RNAs were re-extracted with Nucleospin RNAII (Macherey-Nagel, Düren, Germany) according to the manufacture's instruction. RNA dissolved in 60 μl of DEPC DDW was precipitated by ethanol containing 1 μl of 20 mg/ml glycogen (Wako, Osaka, Japan), and then reconstituted in 10 μl DEPC DDW.

A Super SMART PCR cDNA Synthesis kit (Clontech, Mountain View, CA) was employed for cDNA synthesis according to the manufacture's instruction. Subtractive hybridization was then carried out with a CLONTECH PCR-Select cDNA Subtraction Kit (Clontech) according to the manufacture's instruction, with Hatschek's pit cDNA as the tester and muscle cDNA as the driver. Nucleotide sequences of randomly selected 200 colonies obtained by the subtraction were determined with an ABI 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA), and were analyzed with the ATGC software (Genetyx, Tokyo, Japan). Homology searches were carried out with the Blastx and the Blastn programs

provided by NCBI (http://www.ncbi.nlm.nih.gov/) .

2-2-4 EST analysis of Hatschek's pit dissected by LMD

About 1000 sections which included Hatschek's pit were used for construction of a cDNA library. Total RNA was extracted by the same method as mentioned above, and was dissolved in 10 μl of DEPC DDW. The cDNA library was constructed with a SMART cDNA Library Construction Kit (Clontech) according to the manufacture's instruction. Small PCR products were removed from clones obtained, after the lengths of inserts were checked by colony PCR with the primer sets that were included in the SMART cDNA Library Construction Kit, that is, TriplEx2 5' primer (5'-tccgagatctggacgagc-3') and TriplEx2 3' primer (5'-taatacgactcactataggg-3'). PCR was carried out under following conditions: initial denaturation at 95^{o}C for 4 min, 25 cycles at 95^{o}C for 1 min, 55^{o}C for 1 min, 72^{o}C for 1 min and additional extension at 72^{o}C for 2 min. Clones which included an insert longer than 300 bp were screened by agarose gel electrophoresis, and then nucleotide sequences of 2112 selected clones were analyzed to establish ESTs under contract to Takara Bio Inc.

The amino acid sequences predicted from the nucleotide sequences of the ESTs were then compared with the non-redundant protein database set of NCBI. Meanwhile, a probable presence of proteins was surveyed with the Blastx and a use of the genome data set of *B. floridae* (http://genome.jgi-psf.org/Brafl1/Brafl1.home.html). To link the information from ESTs on the presence of proteins with putative functions, annotation of sequenced data was attempted according to gene ontology terms by use of the program KAAS (KEGG Automatic Annotation Server: http://www.genome.jp/tools/kaas/), which provides functional annotation of genes by Blast comparisons (single best hit) against the manually curated KEGG Genes database (Moriya et al., 2007).

2-2-5 Analysis of micro-proteomics of Hatschek's pit

Tissues of Hatschek's pits from the same number of males and females were collected with fine forceps under a stereoscopic microscope before and after spawning in the breeding season. Tissues from 10 animals were pooled and pestled in a sample buffer (1% SDS; 1% 2-mercaptoethanol; 50 mM Tris-HCl, pH 6.8; 5% glycerol; and 0.01% coomassie brilliant blue), and then heated at 100^{o}C for 5 min. These protein solutions including micro-proteomics of Hatschek's pit were analyzed by SDS-PAGE with a 15% gel which was optimized for Tricine-SDS PAGE (ATTO, Tokyo, Japan). The gels were stained with a silver staining kit (Wako) according to the manufacture's instruction. Several bands shown by the silver staining were cut out by a razor blade.

A LC-MS/MS analysis of proteins and consecutive data analysis were contracted out to APRO life Science Institute, Inc. (Tokushima, Japan). Protein solutions isolated from the gels were blanched and treated with trypsin at 35^{o}C for 20 hours. Afterward, the solutions were analyzed by LC-MS/MS with a Paradigm MS2 (Michrom BioResources, Auburn, CA) HPLC system equipped with a Magic C18 column (Michrom BioResources), followed by an analysis with a Q-Tof2 (Waters Micromass, Milford, MA). Obtained data of fragmented ion spectrum were analyzed with Mascot Search program (Matrix Science, London, UK) against the non-redundant protein database set of NCBI and the predicted amino acid sequences from *B. floridae* downloaded from the Joint Genome Institute.

2-3 Results

2-3-1 Gene expression profile of Hatschek's pit

Sequences obtained by subtractive hybridization between Hatschek's pit and muscle cDNAs are listed in Table 2-1. Sequences of individual clones were automatically

contigated with the EGassemblar provided by the GenomeNet (http://www.genome.jp/ja/), and finally 14 contigs were assembled. A Blastx search against non-redundant amino acid sequences did not show any pituitary hormones in the 14 contigs. Any significant similarities were not found in seven out of 14 contigs.

Clones which included insert sequences longer than 300 bp were selected from constructed cDNA library of Hatschek's pit by colony PCR. Determination of nucleotide sequences in 2112 clones in total resulted in acquisition of sequence data of 2098 ESTs, from which 14 ESTs were removed by failures of sequencing or a shortage of sizes to analyze. The distribution of nucleotide lengths in these clones is shown in Fig. 2-4. The average of nucleotide sequence sizes was 556 bp. Assembly of contigs with the EGassemblar showed that 1641 non-redundant unigene sequences consisted of 170 contigs, which were composed of more than two ESTs, and 1471 singleton ESTs (Table 2-2).

The deduced amino acid sequences were compared with the database of NCBI non-redundant database and the predicted protein data of *B. floridae* by use of the Blastx program. Among 1641 unigene sequences, 1164 sequences (71%) matched with non-redundant data of NCBI database, while 487 sequences (30%) matched with predicted protein data of *B. floridae.* The sequences that matched with high E-value in the NCBI data were hypothetical proteins derived from amphioxus. These sequences matched with particular protein data by Blast search using *B. floridae* genome, although several sequences failed to match despite a significant similarity with amphioxus protein data by the NCBI Blast search. For instance, contig114 and contig20 matched with cytochrome c oxidase subunit I and NADH dehydrogenase subunit 5 with E-value 0.00; however, similar data were not found in the genome of *B. floridae*. Most sequences that matched with low E-value with NCBI data did not match with any data of *B. floridae*.

Possible functions for Hatschek's pit were assigned by use of the KAAS program (Table

2-3). Among 1164 sequences, homologous with that of some protein, approximately 19% were assigned to the KEGG pathways. The remaining 81% were not assigned to any sequences probably due to a lack of EC numbers in the initial Blast analysis or the missing of homology to known pathway genes. Among the categories of cellular processes, six sequences were assigned to five categories of the endocrine system, that is, insulin signaling pathway, SREBP1 (sterol regulatory element binding transcription factor1); PPAR (peroxisome proliferator-activated receptor) signaling pathway, phospholipid transfer protein and ubiquitin C; GnRH signaling pathway, activator protein-1 (AP-1); progenterone-mediated oocyte maturation, heat shock protein 90A; and melanogenesis, frizzled 8.

Concerning those sequences not assigned to the KEGG pathways, all Blast search data were surveyed by eye. As a result, the transcripts related to secretory substances and secretory processes were found in my present data (Table 2-4). As for processing, proprotein convertase (PC) 6B and one type of subtilisin/kexin-like peptidase were found. Six sequences were found to relate to secretory processes. Those related to transport of secretory vesicles are Sec14-like protein and Rab-like protein. Vesicle-associated membrane protein (VAMP)-associated protein A, synaptotagmin, and clathrin-associated adaptor complex AP-1 are related to association of secretory vesicles with cellular membrane. RAB GDP-dissociation inhibitor is a regulator of Rab which is a member of small G-protein and has an important role in secretion. Sequences similar to the enzymes for modification of carbohydrates are mannan endo-1,4-β-mannosidase, α-mannosidase, glycoside hydrolase, and sialyltransferase. Furthermore, similar sequences with rhamnose-binding lectin precursor and mucin-5B known as components of mucus were shown.

2-3-2 Analysis of micro-proteomics of Hatschek's pit

The solutions obtained from dissected and pestled Hatschek's pits were separated by

SDS-PAGE. Significant changes were not detected in band patterns between the tissues collected before and after spawning (Fig. 2-5). Because muscle tissues were likely to be contaminated in Hatschek's pits at the time of dissection, the patterns of SDS-PAGE were compared between Hatschek's pit and the muscle. Then, noticeable bands for Hatschek's pit and neighboring upper and lower areas were cut out and analyzed by LC-MS/MS. In particular, the areas of low molecular weight (lower than 40 k) were focused on for searching peptides related to pituitary hormones. Results of searches with Mascot on the data of LC-MS/MS were listed in Table 2-5. As shown in the list, almost all peptides matched with muscle-related proteins, such as myosin light chain and troponin. Housekeeping molecules, such as a kind of histones were also included. Any pituitary hormones and pituitary related proteins were not found in Hatschek's pit.

2-4 Discussion

The present study showed that gene transcripts and translated products that related to pituitary hormones of vertebrates were not detected in Hatschek's pit, although the analysis of EST indicated expression of genes encoding proteins that are required for transport and release activity of secretory cells.

2-4-1 Deficit of pituitary hormone-related gene transcripts in Hatschek's pit

In my present study on Hatschek's pit, molecular analyses of gene transcripts in combination with the subtractive hybridization and the construction of EST did not show any nucleotide sequences related to so-called pituitary hormones in vertebrates, especially, in gnathostomes.

Tissue samples of Hatschek's pits were collected by the LMD technique which helped accurately collect the target tissues with the least contamination of undesired tissues other

than Hatschek's pits. In contrast to the previous studies on tissue collection from broad areas, this strategy is considered to enable me reliable investigation of the function of Hatschek's pit. Anyway, this is the first approach to the confined investigation of the gene expression of Hatschek's pit itself, and the number of clones that I obtained was reasonable, because the present EST included transcripts of various genes related to activity of secretory cells. Thus, it would be true that pituitary hormone-related genes, if any, are not transcribed, or transcribed at very low levels that could not be sufficiently amplified by the present PCR.

2-4-2 Characterization of transcripts of Hatschek's pit

The analysis with the KAAS program showed expression of several genes related to regulation of synthetic, processing and release activity of secretory cells, such as transcription factors like AP-1, chaperones like HSP90, proprotein convertase like PC6B, and many secretory vesicle-associated proteins. As was described in Introduction, the presence of electron dense granules in glandular cells was reported by Tjoa and Welsch (1974). Sahlin and Olsson (1986) also observed many granules and developed apical mitochondria in Hatschek's pit cells. They further suggested the endocrine function of Hatschek's pit. These previous findings strongly suggest that not all but some cells secrete some proteinaceous or peptidergic chemicals. My present results strongly support the previous idea at molecular levels, although the natures of chemicals remain to be clarified.

On the other hand, the EST analysis showed the presence of precursors for rhamnose-binding lectin and mucin-5B. Lectin is a group of sugar-binding proteins that recognize specific carbohydrate structures and agglutinate in a variety of animal cells by binding to cell-surface glycoproteins and glycolipids. Rhamnose-binding lectins were obtained mostly from various fish eggs (Tateno et al., 1998; Watanabe et al., 2008), and also from the skin mucus of ponyfish (Okamoto et al, 2005). Since gene transcripts for

rhamnose-binding lectin precursor and mucin were detected in Hatschek's pit, it is conceivable that Hatschek's pit secretes mucus substances. On the basis of morphological evidence, many researchers postulated that, in Hatschek's pit, food particles are entangled before being transported posteriad (Franz, 1923, 1927; Barrington, 1963; Welsch and Welsch, 1978; Sahlin and Olsson, 1986, Ruppert, 1997).

2-5 Conclusion

In this chapter, I tried to find homologues of pituitary hormones in Hatschek's pit. Hatschek's pit was successfully isolated by the LMD method. The analyses of gene transcripts and proteins of Hatschek's pit did not detect expression of pituitary hormone-related genes. However, genes related to secretory activity of cells were expressed in Hatschek's pit. Taken together, Hatschek's pit probably secretes some proteinaceous substances like mucus into the oral cavity, although not related to endocrine functions.

Table 2-1. Summary of subractive clones obtained from the cDNA library constructed from Hatschek's pits by subtraction analysis. The titles of columns, nr BlastX indicates the nucleotide homology search by BLAST using the non-redundant protein database DDBJ/EMBL/GeneBank.

Name of unigene	Number of clones	NCBI nr-BlastX	E-value
B2_50	1	protein A [Enterobacteria phage phiX174]	1.00E-93
contig3	3	RNA binding motif protein 42	2.00E-63
contig4	2	hypothetical protein	6.00E-06
B_01	1	dehydrin-like protein	0.013
265	1	hypothetical protein	0.92
671	1	amidophosphoribosyltransferase	1.6
contig1	2	No hit	—
contig2	6	No hit	—
3	1	No hit	—
163	1	No hit	—
480	1	No hit	—
715	1	No hit	—
B2_63	1	No hit	—
B_41	1	No hit	—

Table 2-2. Profiles of EST clones obtained from the cDNA library constructed from Hatschek's pits.

Feature	Value
Number of totall sequences	2098
Average length of sequences	556(bp)
Number of contigs	170
Number of singletons	1471
Number of unique sequences	1641

Table 2-3. Categorization of EST clones using a KEGG software of a biological pathway mapping.

Categories	Number of unigene	%
Metabolism	**67**	**31**
Carbohydrate Metabolism	5	2.3
Energy Metabolism	32	14.8
Lipid Metabolism	6	2.8
Nucleotide Metabolism	6	2.8
Amino Acid Metabolism	6	2.8
Metabolism of Other Amino Acids	4	1.9
Glycan Biosynthesis and Metabolism	3	1.4
Metabolism of Cofactors and Vitamins	4	1.9
Biosynthesis of Secondary Metabolites	1	0.5
Genetic Information Processing	**66**	**30.6**
Transcription	1	0.5
Translation	60	27.8
Folding, Sorting and Degradation	5	2.3
Environmental Information Processing	**20**	**9.3**
Membrane Transport	1	0.5
Signal Transduction	13	6
Signaling Molecules and Interaction	6	2.8
Cellular Processes	**63**	**29.1**
Transport and Catabolism	8	3.7
Cell Motility	4	1.9
Cell Growth and Death	4	1.9
Cell Communication	11	5.1
Circulatory System	11	5.1
Endocrine System	6	2.8
Immune System	15	6.9
Nervous System	2	0.9
Sensory System	1	0.5
Behavior	1	0.5
total	**216**	**100**

Table 2-4. Transcripts related to secretion in EST. The title of NCBI nr means BLAST search without redundunt. Bf ; *Branchiostoma floridae* .

Number of unigene	NCBI nr-BlastX	E-value	Bf genome-BlastX
related to processing			
Contig83	hypothetical protein BRAFLDRAFT_91496 [*B. floridae*]	2E-72	91496| similar to proprotein convertase 6B
25-38_TriplEx25	hypothetical protein BRAFLDRAFT_91496 [*B. floridae*]	3E-13	91496| similar to proprotein convertase 6B
30-46_TriplEx25	hypothetical protein BRAFLDRAFT_91496 [*B. floridae*]	8E-11	91496| similar to proprotein convertase 6B
related to transport			
Contig82	hypothetical protein BRAFLDRAFT_69231 [*B. floridae*]	3E-53	69231| SEC14-like 1
15-21_TriplEx25	hypothetical protein BRAFLDRAFT_56868 [*B. floridae*]	5E-10	246042| rab-like protein
25-42_TriplEx25	hypothetical protein BRAFLDRAFT_68601 [*B. floridae*]	1E-71	68601| VAMP (vesicle-associated membrane protein)-associated protein A
27-7_TriplEx25	hypothetical protein [*Rattus norvegicus*]	6E-12	263902| similar to synaptotagmin
5-38_TriplEx25	rab GDP-dissociation inhibitor [*B. belcheri*]	3E-16	105147| RAB GDP-dissociation inhibitor
13-79_TriplEx25	clathrin-associated adaptor complex AP-1 small chain sigma1 [*B. belcher* i]	5E-32	278617| clathrin-associated adaptor complex AP-1 small chain sigma1
related to glycosylation			
15-19_TriplEx25	hypothetical protein BRAFLDRAFT_98958 [*B. floridae*]	1E-70	98958| Mannan endo-1,4-beta-mannosidase precursor
19-37_TriplEx25	hypothetical protein BRAFLDRAFT_287967 [*B. floridae*]	1E-16	112583| alpha-mannosidase
4-92_TriplEx25	hypothetical protein BRAFLDRAFT_65994 [*B. floridae*]	1E-11	65994| Glycoside hydrolase, family 2, sugar binding
30-78_TriplEx25	hypothetical protein BRAFLDRAFT_249937 [*B. floridae*]	0.003	251148| sialyltransferase 8A isoform 1
related to secretion substance			
Contig43	rhamnose-binding lectin precursor [*B. belcheri*]	1E-47	110271| rhamnose-binding lectin precursor
Contig66	rhamnose-binding lectin precursor [*B. belcheri*]	2E-32	110271| rhamnose-binding lectin precursor
Contig87	rhamnose-binding lectin precursor [*B. belcheri*]	1E-44	110271| rhamnose-binding lectin precursor
Contig124	rhamnose-binding lectin precursor [*B. belcheri*]	3E-33	110271| rhamnose-binding lectin precursor
18-38_TriplEx25	hypothetical protein BRAFLDRAFT_125057 [*B. floridae*]	8E-36	125057| Mucin-5B precursor
27-16_TriplEx25	hypothetical protein BRAFLDRAFT_125057 [*B. floridae*]	3E-26	125057| Mucin-5B precursor
34-16_TriplEx25	hypothetical protein BRAFLDRAFT_125057 [*B. floridae*]	7E-77	125057| Mucin-5B precursor
42-66_TriplEx25	hypothetical protein BRAFLDRAFT_125057 [*B. floridae*]	7E-51	125057| Mucin-5B precursor

Table 2-4.

Table 2-5. Results of a LC-MS/MS analysis for the protein samples, HP-2 to HP11, separated by SDS-PAGE from a homogenate of Hatschek's pits. In each protein sample, the first three data with a higher score are listed. Score more than 53 indicates an identical or significant homology (P>0.05). The title of NCBI nr means BLAST search without redundunt. Bf ; *Branchiostoma floridae* .

Sample name	NCBI nr	Score	Orgenism	Bf predicted protein database	Score
HP-2	sarcoplasmic calcium-binding proteins II, III, V	330	*B. lanceolatum*	sarcoplasmic calcium-binding proteins II,V,VI,VII	287
	muscle LIM protein	175	*B. belcheri*	muscle LIM protein	175
	putative Rab5	140	*Oncometopia nigricans*	histone H2B	148
HP-3	troponin C	284	*B. belcheri*	Myosin, essential light chain	
	Myosin, essential light chain	256	*B. belcheri*	Myosin, regulatory light chain	
	calmodulin	188	*B. floridae*	calmodulin	
HP-5	histone H4	345	*Diprin pini*	histone H4	345
HP-6	sarcoplasmic calcium-binding proteins II, III, V	230	*B. lanceolatum*	sarcoplasmic calcium-binding proteins II, V, VI, and VII	176
	calcium vector protein	189	*B. belcheri*	myosin regulatorr light chain	175
	myosin regulatory light chain	175	*B. belcheri*	calcium vector protein	146
HP-7	myosin regulatory light chain	389	*B. belcheri*	myosin regulatorr light chain	440
	calcium vector protein	246	*B. belcheri*	myosin regulatorr light chain	192
	Myosin, essential light chain	141	*B. floridae*	calcium vector protein	175
HP-8	myosin light chain alkali	829	*B. floridae*	myosin essential light chain	686
	histone H2A.2	141	*Homo sapiens*	histone H2B	223
	putative cyclosporin A-bindinf protein	73	*Picea abies*	calmodulin-like 3	58
HP-9	myosin light chain alkali	135	*B. floridae*	myogloblin	106
	profilin	120	*B. floridae*	myosin essential light chain	73
	hypothetical protein Mfla_2236	61	*Methylobacillus flagellatus KT*	RNA-binding glycine-rich protein-1	65
HP-10	histone cluster1 H4a	260	*Homo sapiens*	histone H4	260
	myosin light chain alkali	100	*B. floridae*	myogloblin	60
	high-molecular weight cobalt-containing nitrile hydrase subunit alpha	100	*Rhodococcus rhodochrous*	protein kinase c inhibitor	55
HP-11	thyoredoxin	209	*B. floridae*	RNA-binding glycine-rich protein-1	116
	high-molecular weight cobalt-containing nitrile hydrase subunit alpha	100	*Rhodococcus rhodochrous*	unknown	114
	cytoplasmic dynein light chain 2	89	*B. floridae*	cytoplasmic light-chain dynein	89

Table 2-5.

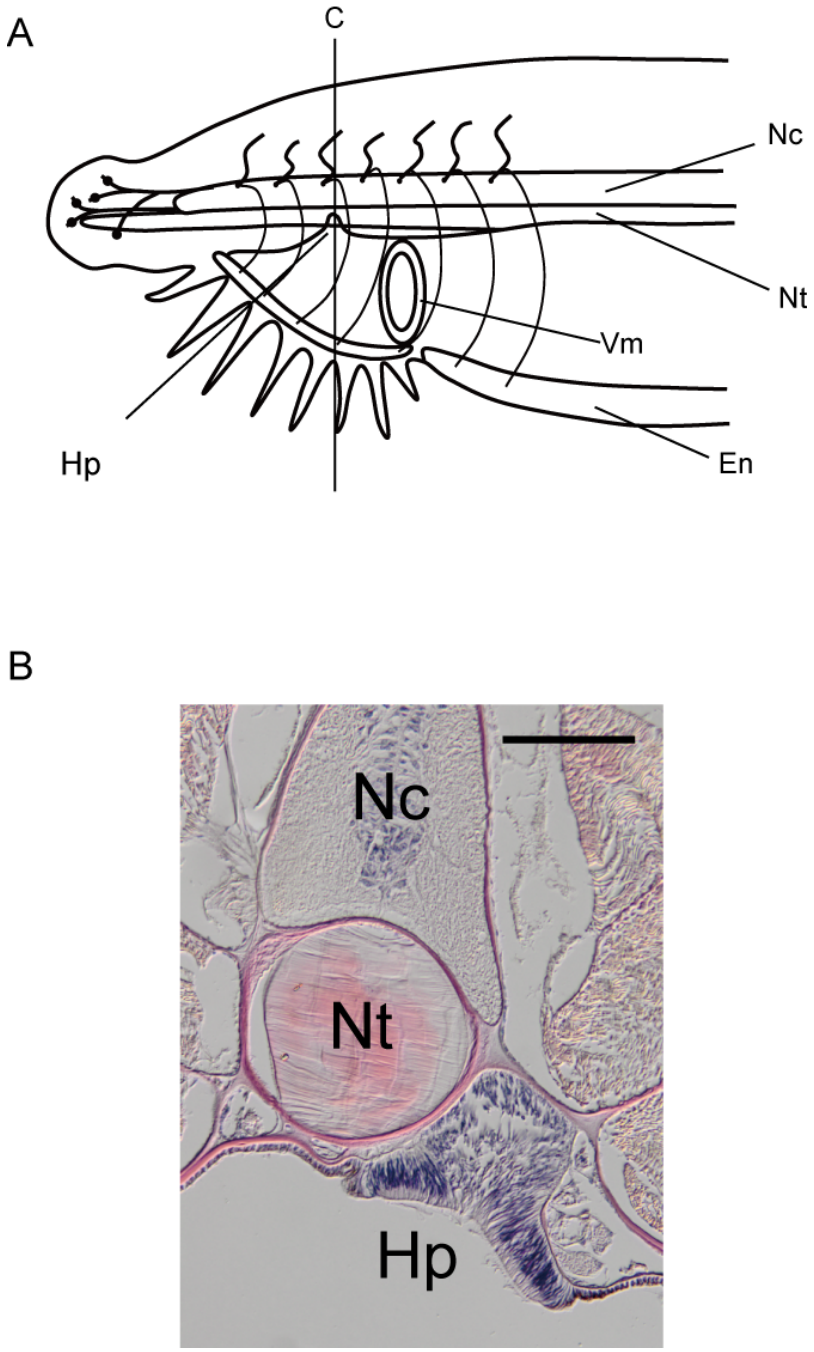

Figure 2-1. Location of Hatschek's pit in the head of amphioxus. (A) Schematic diagram showing relative locations of the nerve cord, magnificent neurons, and Hatschek's pit. (B) Micrograph of the Hatschek's pit region in a hematoxylin and eosin stained transverse section cut at line C in the upper panel. Nc, Nerve cord; Nt, Notochord; Hp, Hatschek's pit; Vm, Velm; En, Endostyle. Scale bar in B, 50 μm.

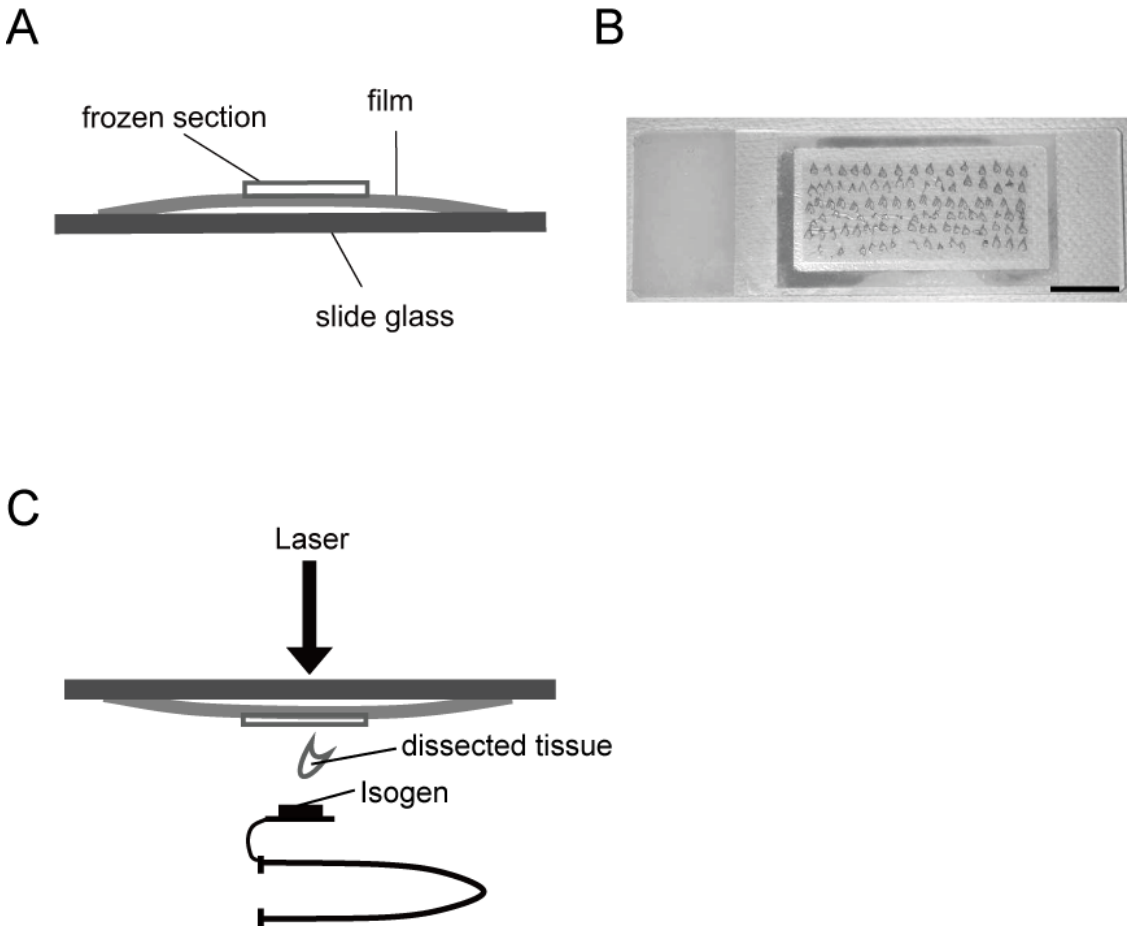

Figure 2-2. Schema of the method to collect small pieces of tissues with a Laser Microdissection system (LMD). (A) A frozen section was cut out from the region of Hatschek's pit, and laid on a plastic film placed on a slide glass. (B) Approximately fifty sections were mounted on a slide glass of 26 mm x 76 mm. Scale; 1cm. (C) The slide glass was inverted and set on a microscope of LMD. The outer marginal part of Hatschek's pit in the section was cut out with a laser beam. The dissected tissue dropped naturally into a cap of microcentrifuge tube set up under the slide glass. The cap was filled with a buffer for isolation of total RNA.

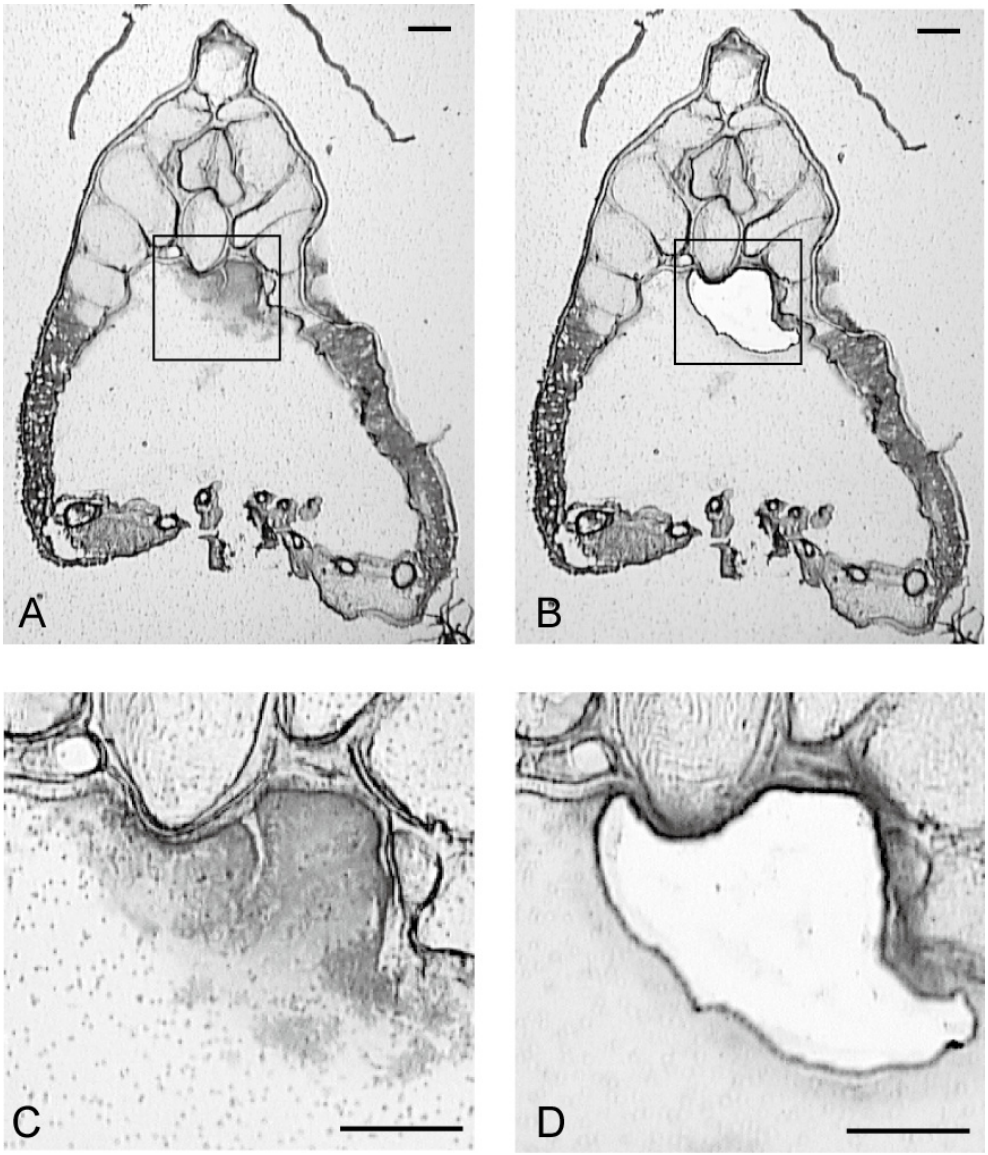

Figure 2-3. Microphotographs of frozen sections before (A, C) and after (B, D) the dissection of a part of Hatschek's pit with a Laser microdissection system (LMD). The area surrounded by the square (A and B) indicates the locus of Hatschek's pit. (C) The region including intact Hatschek's pit to be cut out; and (D) the section from which Hatschek's pit was dissected with LMD. Scale bars, 50μm.

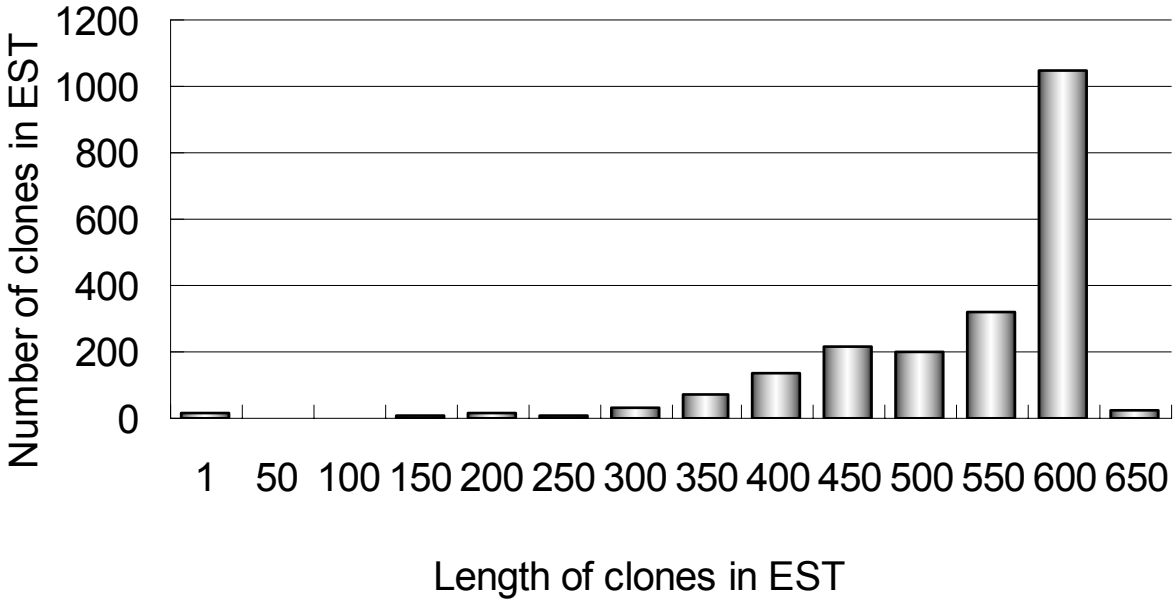

Figure 2-4. Frequencies of EST clones that have particular lengths. The total number of EST clones was 2112. The longest length was 671 bp. EST clones of 600- 649 bp were most abundant. Among the EST clones, 2098 clones were analyzed.

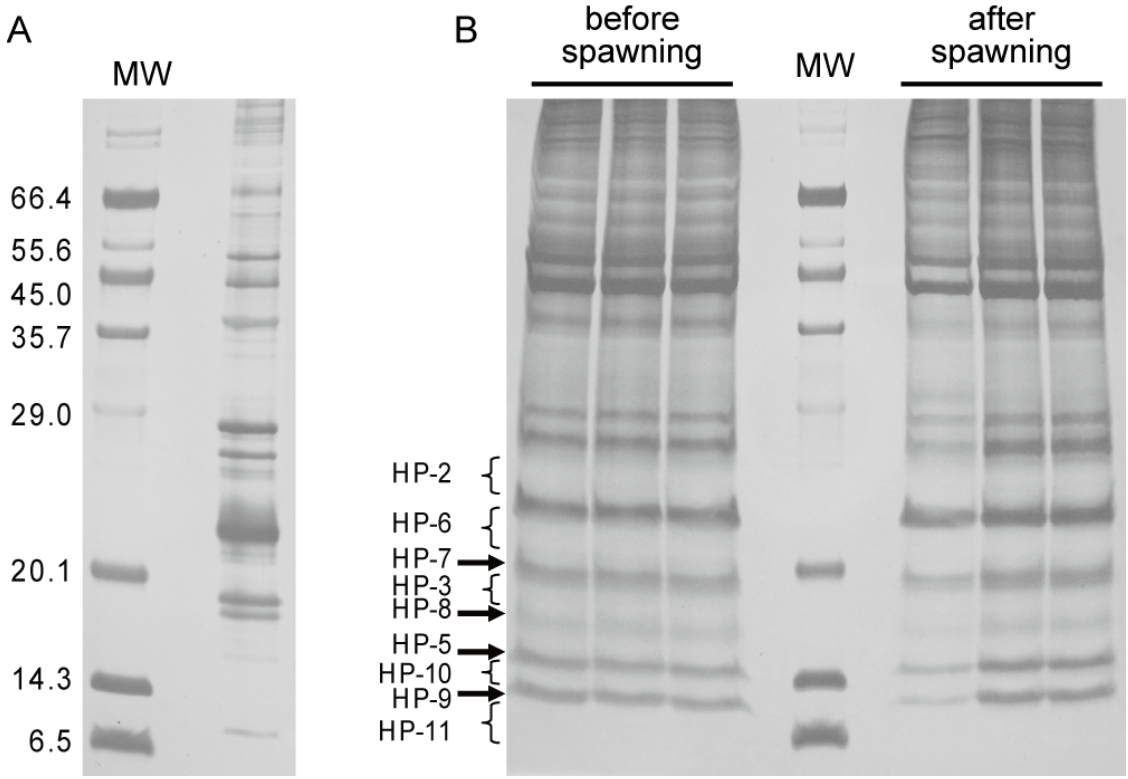

Figure 2-5. Separation of amphioxus proteins from Hatschek's pit and muscle by SDS-PAGE. The bands of SDS-PAGE from homogenized muscles (A) and Hatschek's pits (B) were detected by a silver staining. Hatschek's pits were collected from animals before (three left lanes) and after spawning (three right lanes). MW, a lane of molecular weight marker.

Chapter 3

Structures of Glycoprotein Hormone Subunits in Amphioxus

3-1 Introduction

The pituitary gland secretes seveeral hormones to regulate a variety of physiological processes in vertebrates. On the basis of structural and functional similarity, the pituitary hormones of gnathostomes can be classified into three groups: growth hormone (GH) family, glycoprotein hormone (GPH) family, and proopiomelanocortin (POMC) family. The GH family consists of growth hormone, prolactin (PRL) and somatolactin (SL). They are a single chain polypeptide which has similar structure and gene organization. The GPH family consists of follicle-stimulating hormone (FSH), luteinizing hormone (LH), and thyroid-stimulating hormone (TSH). They are a heterodimeric glycoprotein which is composed of common α subunit and hormone specific β subunit (Li and Ford, 1998; Querat et al., 2000). The POMC family members, adrenocorticotropic hormone (ACTH), melanophore-stimulating hormone, and β-endorphin, are derived from a single precursor (Smith and Funder, 1988; Nakanishi et al., 1979). In spite of diverse habitats and life cycles, all vertebrates have a set of these three hormone families. Meanwhile, most invertebrates do not have homologues of vertebrate pituitary hormones, but POMC is characterized in leech (Salzet et al., 1997) and mussel (Stefano et al., 1999).

Recently published genomic sequences of *B. floridae* provided new genetic information for the endocrine system of amphioxus. According to genome annotation, homologous genes encoding pituitary hormones were not found in the amphioxus genome database using a Blast search with known sequences of vertebrate hormones as a query (Holland et al., 2008). Nevertheless, homologous genes for GPA2 and GPB5 were found in the amphioxus genome (Holland et al., 2008). GPA2 and GPB5 are subunits of recently discovered GPH, thyrostimulin (Nakabayashi et al., 2002). All GPHs are non-covalently associated strong and specific heterodimer which is composed of α and β subunits. These subunits have characteristic cystine knot motifs (Isaacs, 1995; Hearn and Gomme, 2000). The presence of

two α subunits and five β subunits is known in human, where a common glycoprotein hormone α subunit (GPA1) heterodimerize with hormone specific β subunits for FSH, LH, and TSH. In the case of thyrostimulin, a single thyrostimulin-specific α subunit (GPA2) heterodimerizes with a specific β subunit (GPB5), and forms an active hormone molecule (Nakabayashi et al., 2002).

The function of thyrostimulin was investigated in transgenic mice. Overexpression of *GPB5* in mice elevated the serum levels of thyroid hormone, and reduced body weight, despite increased food intake (Macdonald et al., 2005; Okada et al., 2006). Okada et al. (2006) further observed hypertrophy of the thyroid follicles and hyperplasia of the thyroid follicular epithelial cells in *GPB5* overexpression mice. Since thyrostimulin bound TSH receptors *in vitro* and activated them (Nakabayashi et al., 2002), it may function to regulate secretion of thyroid hormone. However, GPA2 and GPB5 subunits of thyrostimulin have distinct characteristics when compared with those of TSH.

Gene transcripts of *GPA2* and *GPB5* were expressed not only in the pituitary, but also in various tissues, such as the brain, pancreas, placenta and gonads (Nakabayashi et al., 2002; Hsu et al., 2002; Okada et al., 2006). In addition, the expression level of *GPA2* was remarkably lower than that of *GPB5* (Nagasaki et al., 2006), and no overt phenotypes were observed in *GPB5* knockout mice (Macdonald et al., 2005). Furthermore, none of the *GPA2* and *GPB5* transcription was induced by thyrotropin-releasing hormone (TRH) in a pituitary cell culture (Nagasaki et al., 2006). These reports raised a question whether thyrostimulin functions physiologically as a heterodimeric hormone, as well as a question whether its biological roles are discrete from those of TSH. Taken together, Nagasaki et al. (2006) pointed out that thyrostimulin is not a part of the typical hypothalamus-pituitary-thyroid axis.

Interestingly, the homologous nucleotide sequences of *GPA2* and *GPB5* are present not only in vertebrates but also in invertebrates (Sudo et al., 2005; Park et al., 2005 Dos Santos et

al., 2009). Since pituitary GPHs, such as FSH, LH and TSH, are found only in vertebrates, pituitary GPHs were considered to be acquired at least in the vertebrate lineage. Therefore, from an evolutional point of view, characterization of amphioxus GPH will provide further understanding about the evolution of vertebrate GPHs. In this chapter, I first reconfirmed whether pituitary hormone genes really do not exist in amphioxus. A homology search and a motif search were applied to the *B. floridae* genome database. Then, I cloned cDNAs encoding thyrostimulin subunits from *B. belcheri*, and predicted structures of deduced proteins. In addition, translated products were characterized by use of recombinant proteins.

3-2 Materials and Methods

3-2-1 Survey of pituitary hormone in the genome database of B. floridae

Blast searches were carried out first to reconfirm the lacks of homologous amphioxus genes encoding PRL, SL, POMC, FSH, LH, and TSH in the genome database of *B. floridae* version 1, followed by resurvey of the version 2. The pattern search approach was then applied to confirm whether GPH subunit genes other than GPA2 and GPB5 subunit genes are really not found in the amphioxus genome. In this approach, I tried to identify the consensus sequence for cystine knot structure, that is, “C(X...)CXGXC(X...)C(X...)CXC” (Fig. 3-1), and N-linked glycosylation site indicated as “NXS or T” in the predicted amino acid sequences from the genome database of *B. floridae* (see Chapter 2 for the download site). Identified protein sequences were further analyzed by confirming a signal peptide, length of amino acids, the number and positions of cysteine residues, and conservation of additional amino acid residues.

3-2-2 Cloning of GPA2 (ampGPA2), GPA2LP (ampGPA2LP) *and* GPB5 (ampGPB5)

Total RNA was extracted from a whole body of single animal as described by Mizuta and Kubokawa (2007), dissolved in 50 μl of DEPC water, and stored at -80°C until use. First-strand cDNA was constructed from total RNA, 1 μg/20 μl of RT reaction, with a PrimeScript™ RNA PCR kit (TaKaRa) with Oligo-dT primer. RT reaction was performed at 45°C for 30 min, 55°C for 10 min, 65°C for 10 min, and 70°C for 10 min, and finally PrimeScript reverse transcriptase was denaturated at 95°C for 5 min. RT products were stored at -30°C until use. PCR of RT products was performed using 2 μl of first-strand cDNA as a template in 20 μl of PCR mixture. The primers used for cloning of *ampGPA2*, *ampGPA2LP*, and *ampGPB5* are listed in Table 3-1. Partial cDNA fragments of *ampGPA2*, *ampGPA2LP*, and *ampGPB5* were obtained from whole body cDNA of amphioxus under following conditions: initial denaturation at 95°C for 1 min, 40 cycles at 95°C for 10 sec, 60°C for 10 sec, 72°C for 30 sec, and additional extension at 72°C for 3 min. The PCR product was applied to an agarose gel electrophoresis. A fragment of expected size was cut out, subcloned into a pCR4 TOPO plasmid vector (Invitrogen, Carlsbad, CA), and sequenced. Full length cloning of *ampGPA2*, *ampGPA2LP*, and *ampGPB5* was carried out following Mizuta and Kubokawa (2007).

3-2-3 Alignment among GPAs and GPBs of various animals

Amino acid sequences of GPAs and GPBs from both invertebrates and vertebrates were aligned with the Clustal W program (Thompson et al., 1994). The DDBJ/GenBank/EMBL accession numbers for amino acid sequences used for the analysis were as follows: human GPA2 (AF260739), takifugu GPA2 (Q4S0U2), fly GPA2 (AY940435), nematoda GPA2 (BN001246), sea hare GPA2 (BN001237), human GPB5 (AF403430), zebrafish GPB5 (XM001343401), fly GPB5 (AF403389), nematoda GPB5 (AF403389), and sea hare GPB5 (AY928334). Amino acid sequences of frog GPA2 (protein ID, 360710) and GPB5 (ID,

379795) were excerpted from the genome database of *Xenopus tropicalis* (http://genome.jgi-psf.org/Xentr4/Xentr4.home.html).

3-2-4 Modeling of 3D structure

The model of three dimensional (3D) structures of ampGPB5 and human GPB5 were constructed by use of the protein modeling routine available at the SWISS-MODEL server (http://swissmodel.expasy.org/) conferring the experimentally determined structures of human FSHβ (hFSHβ, Protein Data Bank, accession code 1XWD, Chain B), human chorionic gonadotropin β (hCGβ, Protein Data Bank, accession code 1HRP, Chain B) and human glycoprotein hormone α (hGPA1, Protein Data Bank, accession code 1HRP, Chain A) as templates. ChemBio 3D Ultra version 11.0 software (Cambridge Soft, Massachusetts, USA) was used to visualize the 3D structure.

3-2-5 Construction of recombinant proteins (Fig. 3-2)

Baculovirus transfer plasmids were constructed by a pFastBac system (Invitrogen) using pDEST10 containing polyhedrin promoter. The cDNAs for *ampGPA2*, *ampGPA2LP* and *ampGPB5*, to which a sequence for 6x His were conjugated at the 3' end, in the entry plasmids were introduced into pDEST10 by a LR recombination reaction. After the reaction, the resulting baculovirus transfer plasmid, such as pDEST10/ampGPA2, ampGPA2LP or ampGPB5, was transformed into *E. coli* BmDH10Bac cells containing a BmNPV bacmid (Motohashi et al., 2005). Through *in vivo* transposition mediated by Tn7 transposase, the reporter cDNAs were transferred into a mini-attTn7 target site of baculovirus shuttle vector (Bacmid) in *E. coli* cells.

After purification, the recombinant Bacmid DNAs were transfected into the *Bombix mori* e21 *(Bme21*) cells to generate recombinant baculoviruses. Three days after transfection,

the culture medium was collected, and the infection was repeated two times for preparation of a high titer virus stock. *B. mori* d17 strain used in the present study was provided by the Institute of Genetic Resources, Kyushu University Graduate School. Silkworm larvae on day 2 of the 5th instar were infected with 2×10^4 pfu of recombinant BmNPVs. The larval legs were cut 5 days after viral injection, and hemolymph was collected from each larva. Hemocytes were then removed by centrifugation at 2,000g for 5 min, and the supernatant was subjected to purification. The supernatant was applied to the column (Bio-Rad Laboratories, Hercules, CA) filled with Ni sepharose resin (GE Healthcare, Tokyo, Japan), eluted by gradually increased immidazole (20-300mM) in elution buffer (respective concentrations of imidazole, 20 mM phosphate, 500 mM NaCl, and 10% glycerol). Fractions were checked by SDS-PAGE with Coomassie Brilliant Blue (CBB) staining, and then further purified by reverse-phase high performance liquid chromatography (RP-HPLC) on a SenshuPak Pegasil ODS (4.6 mm i.d. x 250 mm, Senshu Kagaku, Tokyo, Japan) with linear gradient of 0–60% acetonitrile in 0.05% trifluoroacetic acid. N-terminal amino acid sequences of purified proteins were determined by Edman degradation with a protein sequencer (491 cLC, Applied Biosystems, Foster City, CA) in the pulsed-liquid mode.

3-2-6 Western blot

Western-blotting analysis was performed using anti-6x His antibody (1:5000, Thermo Fisher Scientific, Waltham, MA), rabbit polyclonal antisera against ampGPA2 (1:100) and ampGPB5 (1:200). Laemmli's SDS-PAGE sample buffer including 10% 2-mercaptoethanol was added to a protein solution, heated at 50^oC for 5 min, and separated by SDS-PAGE using 14% polyacrylamide gel. Separated proteins were transferred to a PVDF membrane (ATTO). After a treatment with blocking buffer (1% casein and 1% BSA in phosphate buffered saline, pH 7.4/ 0.1% Tween 20 (PBT)), the membranes were incubated with the first antibody or

antiserum at room temperature for 1 hr. Specific signals were detected with a Vectastain ABC elite rabbit IgG Kit (Vector Laboratories, Burlingame, CA) according to the manufacture's instruction.

3-2-7 Treatment with glycosidase

Each subunit was treated with glycosidase using an Enzymatic Protein Deglycosylation Kit (Sigma-Aldrich, St Louis, MO) according to the manufacture's instruction. Briefly, proteins were denatured by heating at 100^{o}C for 5 min under reducing condition prior to the treatment with glycosidase, and then N-glycosidase F (PNGaseF) was added and incubated at 37^{o}C for 3 hrs. Reaction products were detected by SDS-PAGE and Western blot.

3-2-8 Cross-link analysis

Purified recombinant GPA2, GPA2LP and GPB5 were cross-linked with 1 mM of disuccinimidyl suberate (DSS, Wako) at 25^{o}C for 30 min, and the reaction was terminated by an addition of 1 M Tris-HCl, pH 7.4. Reaction products were detected by SDS-PAGE and Western blot as was mentioned above.

3-3 Results

3-3-1 Presence of homologous genes of vertebrate GPH

Three homologues of GPH subunit genes, two α subunit genes which is homologous to *GPA2* and one β subunit gene which is homologous to *GPB5*, were found by the Blast search against the genome database of *B. floridae* V.1 and V.2, as demonstrated by Holland et al. (2008). Forty-five amino acid sequences of predicted proteins, which included the cystine knot motif, were found after the pattern search approach. Abovementioned three GPH

subunits were included in these 45 sequences. However, only these three homologues satisfied the criteria of GPH subunit in terms of the presence of signal peptide, protein size, the number and additional cysteines, and conservation of additional amino acid residues.

3-3-2 Amphioxus GPH subunit genes

Two genes homologous to *GPA2* of vertebrates were found in the genome database of *B. floridae*. On the basis of their sequences, full-length cDNAs for these two genes (*ampGPA2* and *ampGPA2LP*) were obtained from *B. belcheri* using RT-PCR and RACE. One of the cDNA encoding ampGPA2 is 1205 bp in length, and contains an open reading frame of 378 bp (126 amino acids) (Fig. 3-3). Similarities of amino acid and nucleotide sequences for GPA2 between *B. belcheri* and *B. floridae* (protein ID, 117901) are 93% and 91%, respectively. A putative N-linked glycosylation site was found at Asn97. The other cDNA which encodes ampGPA2LP is 1349 bp in length and contained an open reading frame of 465 bp (155 amino acids) (Fig. 3-4). Similarities of amino acid and nucleotide sequences for ampGPA2LP with the *B. floridae* equivalent (protein ID, 63816) are 82% and 77%, respectively. However, the predicted amino acid sequence does not contain any consensus site for N-linked glycosylation. In terms of the amino acid sequences, the amphioxus GPA2 sequence shares 21.6% and 39.9% identity with human GPA1 and GPA2, respectively, while GPA2LP shares 19.1% and 27.5% identity with human GPA1 and GPA2, respectively.

The length of cDNA encoding ampGPB5 is 1129 bp with an open reading frame of 399 bp (133 amino acids) (Fig. 3-5). The identities of amino acid residues and nucleotides of the coding region for GPB5 are 93% and 94%, respectively, between *B. floridae* and *B. belcheri*. GPB5 shows 39.5% amino acid identity with human GPB5, while 26.8% with hLHβ, 25.2% with hFSHβ, and 24.8% with hTSHβ.

An alignment of amino acid sequences among human, frog, takifugu, fly, nematode, sea

hare, ampGPA2 and ampGPA2LP (Fig. 3-6A) indicates that ampGPA2 contains conserved cysteine residues whose positions are consistent with five S-S bonds in the mature protein. Three S-S bonds contribute to the cystine knot motif (Vitt et al., 2001). When compared with ampGPA2, ampGPA2LP has longer N-terminal amino acid chains. The number and position of N-linked glycosylation sites are not conserved at all in vertebrates. One predicted N-linked glycosylation site at Asn97 of ampGPA2 is located within the finger2 region of the cystine knot motif like GPA2 in fly, whereas vertebrate GPA2s have two glycosylation sites, one in the finger1 region and the other in the heel region of the cystine knot motif (Vitt et al., 2001).

Amino acid residues of GPB5 were aligned among human, frog, zebrafish, fly, nematode, sea hare, and amphioxus (Fig. 3-6B). The cysteine residues necessary for the cystine knot motif are conserved in ampGPB5. One potential N-linked glycosylation site is at Asn121 in the C-terminal region which is different from the cases of vertebrates. GPB5s of fly, nematode and sea hare do not have any glycosylation sites. In vertebrates, one N-linked glycosylation site is located in the finger 2 region.

3-3-3 Modeling of 3D structure of ampGPB5 protein

On the basis of the crystal structure of hFSHβ, a putative model of tertiary structure of GPB5 was completely generated (Fig. 3-7). In this model, the cystine knot motif is formed by three disulfide bridges in ampGPB5 between positions at Cys37–Cys86, Cys62–Cys117, Cys66–Cys119, similar to human GPB5 and hFSHβ. As in human GPB5, the C-terminus of amphioxus GPB5 lacks the last cysteine residue that is crucial to form the disulfide bond (Cys38–Cys122) in the seat belt region of hFSHβ (Fan and Hendrickson, 2005).

When a model of 3D structure of GPB5 was attempted to be constructed by use of hCGβ as a template, as was the usage in human GPB5 (Hsu et al., 2002), modeling was not

successfully achieved for ampGPB5. Cys37, which forms the disulfide-bond necessary for the cystine knot motif, was not aligned to the corresponding cysteine in hCGβ by use of the SWISS-MODEL software. When aa32–36 (DSSLG) of ampGPB5 was replaced by aa25–35 (ASSGNLRTFVG) of human GPB5, then modeling was successfully accomplished. To address whether ampGPA2 and ampGPA2LP form the typical cystine knot motif, the SWISS-MODEL protein folding program was employed by conferring the experimentally determined structure of human GPA1 as a template. However, they could not form a complete cystine knot motif.

3-3-4 Characterization of recombinant GPA2, GPA2LP and GPB5 subunits

The silk worm expression system is capable of N-linked oligosaccharides modification with expressed proteins (Kato et al., 2009). The amino acid sequences of the N-terminals of purified recombinant ampGPA2, ampGPA2LP and ampGPB5 were determined by the Edman degradation sequence method (Figs. 3-8, 3-9, 3-10). Since the results indicated that GPA2 was started from Ala24 (Fig. 3-3), the molecular weight (MW) of recombinant ampGPA2 conjugated with 6x His at its C-terminus was predicted to be 10.8k, while ampGPA2LP was started from Thr26 (Fig. 3-4) and calculated as 14.1k. AmpGPB5, which was started from Ser34 (Fig. 3-5), was predicted to be 11.3k.

The recombinant ampGPA2, ampGPA2LP, and ampGPB5 proteins in the extracts of the expression system were detected with anti-6x His antibody and subunit specific antisera against ampGPA2 and ampGPB5 (Fig. 3-11). All of anti-6x His and subunit specific antisera recognized each band smaller than MW of 14k, although larger than the predicted value. The band of ampGPA2LP seen with anti-6x His was larger than 14.3k. Two or three bands in the same lane were recognized as to recombinant ampGPA2 and ampGPB5. After the treatment with PNGaseF, the upper band of ampGPA2 disappeared and became a single

band (Fig. 3-12A). The upper band of ampGPB5 was diminished after the enzyme treatment (Fig. 3-12B). These results were concordant with the idea that sugars actually occupy the putative N-linked glycosylation sites in the amino acid sequences of ampGPA2 and ampGPB5 expressed in silk worm.

Since dimerization of α and β subunits is necessary for biological activity of GPHs, I tested whether the recombinant subunits could dimerize by cross-linking analysis using a chemical crosslinker, DSS. Incubation with DSS of a mixture of ampGPA2 and ampGPB5, and ampGPA2LP and ampGPB5 showed a presence of high MW band near 29k by Western blotting with anti-6x His antiserum (Fig. 3-13A). The MWs of these bands were almost the same with those estimated from the MW of individual subunits, whereas a mixture of ampGPA2LP and ampGPA2 formed a dimer whose MW was slightly smaller than the predicted value. When the estimated sizes of bands were compared, the MW of band from the mixture of ampGPA2LP and ampGPB5 was larger than that from the mixture of ampGPA2 and ampGPB5. Anti-ampGPA2 and anti-ampGPB5 antisera also showed the presence of high MW bands near 29k, which was similar to the predicted size of the heterodimer (Fig. 3-13B).

After the incubation of single subunits with DSS, Western blotting with anti-6x His antibody detected a presence of high MW band near the 29 k marker in each of ampGPA2LP and ampGPB5 protein solutions; however, ampGPA2 solution did not show any high MW bands (Fig. 3-14A). The MW of high molecular weight band was about 2-fold that obtained by ampGPB5 alone. When ampGPB5 specific antiserum was used, it also showed a similar MW band (Fig. 3-14B). AmpGPA2 specific antiserum, this time, detected a high MW band in ampGPA2 solution, although the reaction was faint and the size was smaller than that of ampGPB5 (Fig. 3-14B).

3-4 Discussion

My approach in this chapter showed that amphioxus has thyrostimulin, a family member of GPH, but does not have any pituitary hormone homologues. Thyrostimulin is composed of α and β subunits, named ampGPA2 and ampGPB5, both of which share a similar cystine knot motif and one N-linked glycosylation site. Glycosylation at this site was confirmed by a treatment of recombinant proteins with glycosidase. Chemical cross-linking analysis showed that ampGPA2 and ampGPB5 subunits can form homo- and heterodimmer.

3-4-1 Structure of amphioxus ampGPA2, ampGPA2*LP, and ampGPB5*

Regardless of extensive survey, homologues of pituitary hormone genes were not found in the genome database of *B. floridae*. However, homologous nucleotide sequences encoding thyrostimulin GPA2 and GPB5 subunits were found, so that I cloned and sequence-analyzed cDNAs encoding these subunits from *B. belcheri*. The proteins predicted from nucleotide sequences were referred to as ampGPA2 and ampGPA2LP, and ampGPB5. ampGPA2LP does not have a N-linked glycosylation site.

Multiple alignments showed that all three subunits have cysteine residues necessary for formation of the cystine knot motif. Comparison of 3D models showed that this motif is highly conserved in some GPHs, such as putative 3D model of ampGPB5, human GPB5 and hFSHβ which was used for modeling of GPB5 as a template, but not in hCGβ. The analysis with ampGPB5 chimera indicated that amino acid sequences of the N-terminal, particularly those in the adjacent upstream of the first cysteine, are important for organization of the 3D structure of cystine knot structure. Comparison of amino acid sequences in this region (Fig. 3-15) supports this idea and further suggests that the amino acid residues followed by the first

cysteine region of ampGPB5 are possibly more closely related to hFSHβ than hCGβ.

My present study showed that, like subunits of vertebrate thyrostimulin, ampGPA2 and ampGPB5 share one N-linked glycosylation site, in contrast to a lack of this site in invertebrate thyrostimulin and ampGPA2LP. A combination of a treatment with glycosidase and Western blotting with subunit specific antisera displayed that the glycosylation sites could be actually glycosylated by the silkworm recombinant expression system, although the nature of sugars was remained to be clarified. Since oligosaccharides of human GPA2 function to activate receptors (Okajima et al., 2008), N-linked glycosylation of ampGPA2 and ampGPB5 may be important for binding and/or activation of receptors of thyrostimulin.

3-4-2 Formations of amphioxus GPH

As mentioned previously, dimerization as well as glycosylation of α and β subunits is generally important for biological activity of GPHs, so that I attempted to clarify whether ampGPA2 and ampGPB5 form a heterodimer. Chemical cross-linking of recombinant subunits showed that they formed homodimer in addition to heterodimer.

When individually treated with DSS, all of recombinant ampGPA2, ampGPA2LP and ampGPB5 showed the presence of high MW bands, suggesting that they formed homodimer by chemical cross-linking. Although the use of anti-6x His antibody failed to demonstrate such a band in reaction product of ampGPA2, it was probably due to masking of the 6x His epitope tags at the C-terminus by a portion of dimerized protein. A question arising here is whether such homodimerizarion really occurs *in vivo*.

Chemical cross-linking of a mixture of two different subunits yielded heterodimeric products in all combinations, although the amounts of products were not the same. A problem was that chemical cross-linking produced homodimers. However, since the use of subunit specific antisera showed that the high MW band in a mixture of ampGPA2 and

ampGPB5 was larger than those in single ampGPA2 and ampGPB5, I considered that cross-linking really produced a heterodimer. Taken together, the present results evidenced that amphioxus GPH subunits at least have potential to dimerize in *in vitro* condition.

My present cross-link analysis showed much of monomers remained. Interestingly, there were a few lines of evidence for the function of GPH subunit as monomer or homodimer since the 1980's (Begeot et al., 1984). Ectopic free β subunit of hCG may bind and activate a component of TGFβ receptor (Iles, 2007). Free α subunit generated in the endometorial stromal cells was involved in the regulation of endometorial cell differentiation (Nemansky et al., 1998). In addition, αα homodimer of hCG is secreted from cultured carcinoma cells (Krause et al., 2007). These reports remind me that one of important questions in understanding of the evolution of GPH is which is the ancestral functional form, monomer, homodimer or heterodimer. To clarify this question, it is indispensable to obtain native ampGPA2, ampGPA2LP, and ampGPB5.

3-5 Conclusion

I tried to find homologues of pituitary hormones in Hatschek's pit. Hatschek's pit was successfully isolated by the LMD method. The analyses of gene transcripts and proteins of Hatschek's pit did not detect expression of pituitary hormone-related genes. Amino acid sequences of α and β subunits, ampGPA2, ampGPA2LP and ampGPB5, were predicted from nucleotide sequences of *B. belcheri* cDNAs. They share cysteine residues for the cystine knot motif. The amino acid sequence of ampGPA2LP does not include any glycosylation sites, and has the longer N-terminal when compared with GPA2 of amphioxus and other animals, whereas ampGPA2 and ampGPB5 have one N-linked glycosylation site. Glycosylation at this site was confirmed by a treatment of recombinant proteins with glycosidase. Cross-link analysis further showed that homo- and heterodimer of subunits

were formed by chemical cross-linking. These results indicate that amphioxus thyrostimulin subunits are glycosylated, and able to dimerize.

Table 3-1. Primer sequences used for cloning of *ampGPA2*, *ampGPA2LP*, and *ampGPB5*.

Primer name	Sequence (5' to 3')	Experiment
GPA2 partial-F	TAGAGGCTACTGTGAGTCCATA	partial fragment of GPA2
GPA2 partial-R	GGTTATCGTCGCAGATGCTACA	partial fragment of GPA2
GPA2LP partial-F	GCGCTATATGCCAATGTTAGCCGTAC	partial fragment of GPA2LP
GPA2LP partial-R	CAGCAGACTCGATAGTGTAGGTC	partial fragment of GPA2LP
GPB5 partial-F	CTGGGCTGCGACGTCTGGAGAG	partial fragment of GPB5
GPB5 partial-R	ATACTAACGTTTCGCAGGAC	partial fragment of GPB5
GPA2GSP-R	CGCAGATGCTACAGGCGCAGCTGGAT	5' RACE first PCR of GPA2
GPA2GSP-R_Nest	GCGCAGCTGGATGCCGAGTAGAGGGTTT	5' RACE nested PCR of GPA2
GPA2GSP-F	GCAGGGGTCGAGCGGGAACCACGTCAT	3' RACE first PCR of GPA2
GPA2GSP-F_Nest	AGCCTGCTGTGACATCGCCTCCACACAT	3' RACE nested PCR of GPA2
GPA2LPGSP-R	GCAGGCGTTGATGAGAACCGTAG	5' RACE first PCR of GPA2LP
GPA2LPGSP-R_Nest	GCCAGATCCCACTCAGACGTCAG	5' RACE nested PCR of GPA2LP
GPA2LPGSP-F	GATGAACGGGAGCAGAACGAAATAGG	3' RACE first PCR of GPA2LP
GPA2LPGSP-F_Nest	CAGACCTCCCTGTCCTCTTACGC	3' RACE nested PCR of GPA2LP
GPB5GSP-R	GGGTCGACCCCTGGCTGGCAGTTCT	5' RACE first PCR of GPB5
GPB5GSP-R_Nest	TTGTCCTGGTCTCCACCCGGTCGTACGT	5' RACE nested PCR of GPB5
GPB5GSP-F	CCGAGAAGGAGGGATGCGAGCGGCTACA	3' RACE first PCR of GPA2
GPB5GSP-F_Nest	GCGTAGACGCTTGCAAGGGCCGCTGTG	3' RACE nested PCR of GPA2

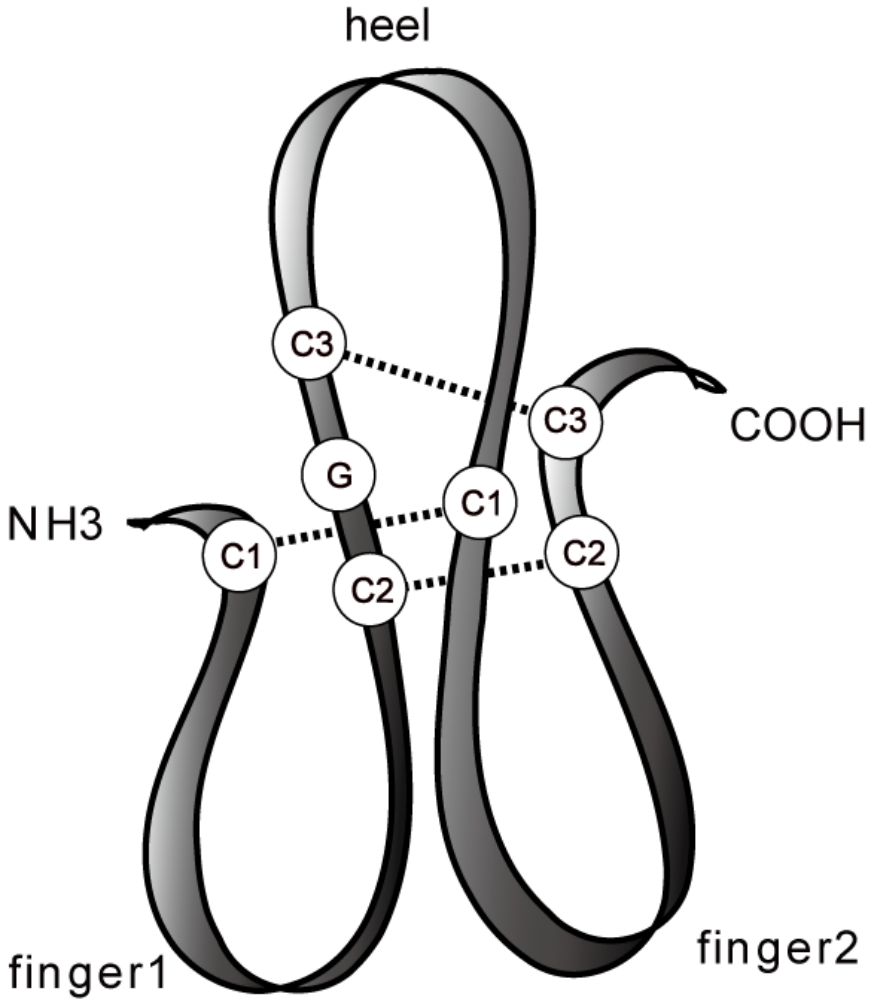

Figure 3-1. Schematic drawing of the cystine knot structure. Dotted lines connecting cysteine residues indicate disulfide bonds. Three disulfide bonds contribute to the cystine knot structure. C, cysteine; G, glycine.

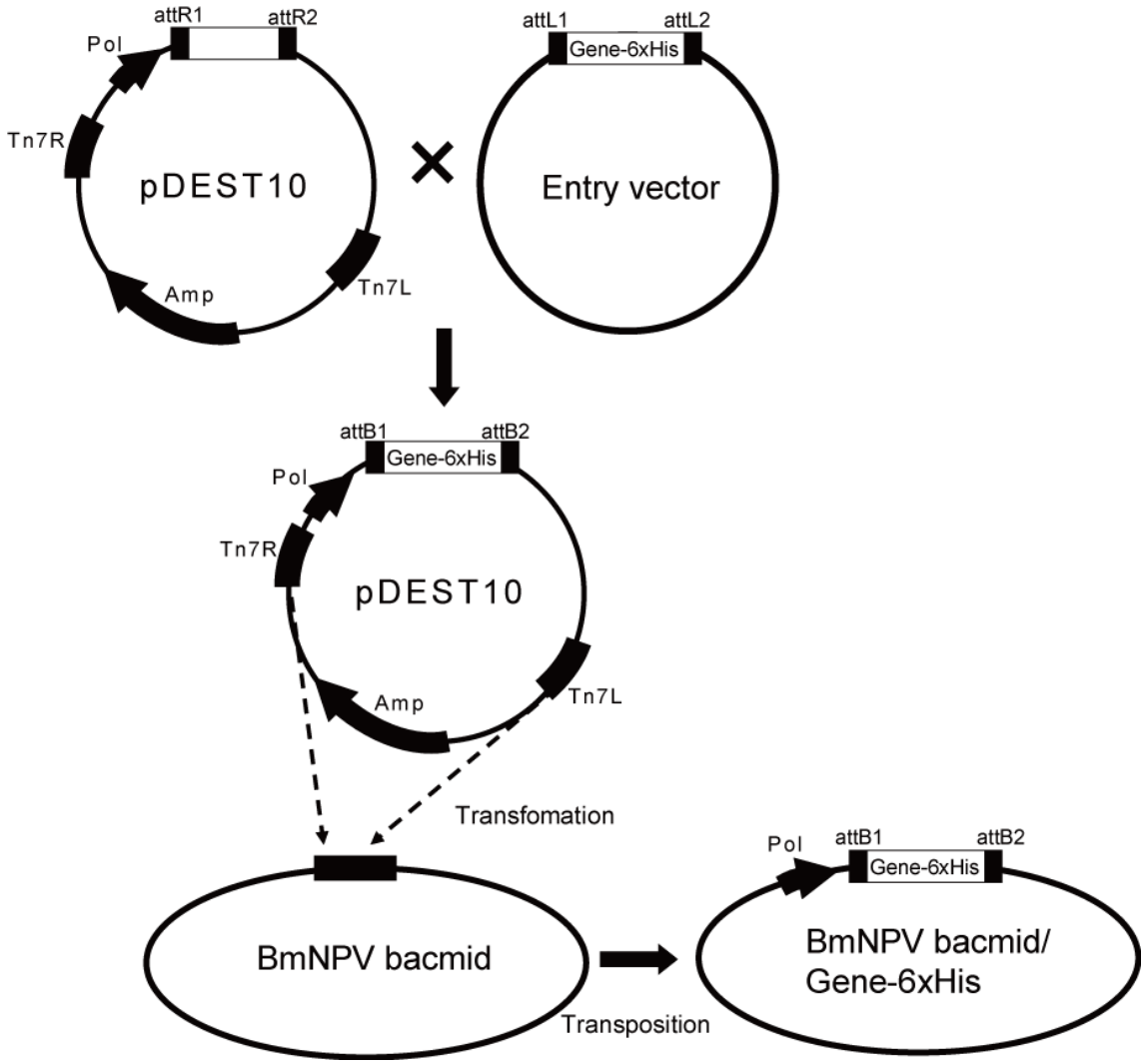

Figure 3-2. Schema of the procedure for expression and construction of recombinant ampGPA2, ampGPA2LP, and ampGPB5 with a BmNPV bacmid system. Each cDNA for polypeptide conjugated with 6x histidines at the C-terminal was introduced into a vector pDEST10 by LR recombination reaction. The vectors were transformed into *E. coli* BmDH10Bac cells containing a BmNPV bacmid. Recombinant DNA obtained through these processes was designated as BmNPVbacmid/Gene-6x His. Amp, ampicilin resistant open reading frame. attB1, attB2, attL1, attL2, attR1, and attR2, recombination sites. Pol, polyhedrin promoter. Tn7L, Tn7 left arm. Tn7R, Tn7 right arm.

```
   1 TATATACCGACCACCTTAAGCAATCACGACAAGAGGTGCGGTGACAAAAAGGCCAGTTAA  60
  61 TGATTGACAGAGAGGCCGACCTGACGGAAGGGCCCGAGGGTTGGACGTCAGCAATCGTGA 120
 121 GCGTGCATCATCTCCACAGCGGGAATCCTGCGTTTCCCCAGACCTGCCTGTCCCCTGCTG 180
 181 GGCTGTCAGCCGTTGCTAACAAACTGTGAAGTTGGAAACAAATTGCTGACCGGGGGATGG 240
 241 TATAAGATAGGATCCATGTAGACTACATTTCGCTTATCGCATCCGGCCGCCGATGAGTAG 300
 301 CACTTAGGGGAGCATTCAGCAGGAATCATAGAACATTCAGAAGCCAGTCTGGGAGAAACA 360
 361 GGCACGATGCAGAGACTACTGTCGTGGGTACTGCTGCTGACGATCCTTCTGTCTTCATCC 420
   1       M  Q  R  L  L  S  W  V  L  L  L  T  I  L  L  S  S  S   18

 421 GACATGGGCAGGACGCAGGCCCCCTGGTACAGGCCCGGCTGTCATCTCGTAGGAGTTGAC 480
  19 D  M  G  R  T  Q  A  P  W  Y  R  P  G  C  H  L  V  G  V  D   38

 481 AAATCTGTGGAGGTGCCTGGGTGCCAGAGACAGACTGTCCGAGTGAACGCCTGTAGAGGC 540
  39 K  S  V  E  V  P  G  C  Q  R  Q  T  V  R  V  N  A  C  R  G   58

 541 TATTGTGAGTCCATAGCCTTCCCGTCCTCCAGCACCACACGGCAGGGGTCGAGCGGGAAC 600
  59 Y  C  E  S  I  A  F  P  S  S  S  T  T  R  Q  G  S  S  G  N   78

 601 CACGTCATCACGTCCAGAGCAGCCTGCTGTGACATCGCCTCCACACATGTGGTGAACTTC 660
  79 H  V  I  T  S  R  A  A  C  C  D  I  A  S  T  H  V  V (N) F   98

 661 TCTCTGCGCTGTGGAAATCTGCTGGTCCCCAAAACCCTCTACTCGGCATCCAGCTGCGCC 720
  99 S  L  R  C  G  N  L  L  V  P  K  T  L  Y  S  A  S  S  C  A  118

 721 TGTAGTATCTGCGACGATAACCCGTGACGTCATAGCTTGGCAAGAAGAACCTTGAAGAAG 780
 119 C  S  I  C  D  D  N  P  *                                    126

 781 AGGCTTGTAACCTTTATGGTCTAAAGATATGCTTCTGGTGACGACAGGATTTTTTTTCAA 840
 841 AGATAATGAATGACTTTGAAATAAAGATTTGCTCACCGTGATCACAAAGTGTGAGTTACA 900
 901 GAAAAATCTTGTAGTGTAATGTAGACGGTAAATAAATGCGGCAACAGAATTAAAAGAGCT 960
 961 GTACAGGCCATTTCTGCTTTTGCCTTGATGGCAACTATAATGAGAATAGTAGACATAAAA 1020
1021 ACACAACAGAGACAAGATGCGAATGCACGTACTTGTTGTACTCCCGATGATAAAGCATGT 1080
1081 CGTCCTATGCTGTACAGGTGTTAGCCGTAAGGATTAAAACGAGGACATTTGCAAAAGCTT 1140
1141 CGTCCAACAACCTCATGCAGGAAATAAAGATTATACGCTGAAAAAAAAAAAAAAAAAAAA 1200
1201 AAAAA                                                         1205
```

Figure 3-3. Nucleotide and amino acid sequences of ampGPA2. The underline indicates the signal peptide region. The potential N-linked glycosylation site is circled.

```
   1 GGACACTGACATGGACTGAAGGAGTAGAAAAAAGTGATTCCATGTATAGCTCCATAACAC  60
  61 AAGGATATATTGTATCACCCAATTGACATTTAAAAACAATTTCCTTATCACGACAGACAA 120
 121 ACACAGATTCGTACCAGAATGGACTAAGTAAACTAGGGCGTTGTAATAAAAGGCACTGCT 180
 181 CAAACGTTGTAAAATACATTGAAAATCTCTGCAGCCAACACATCATTTTCACTGAAAAAC 240
 241 AAGTGCGAAAAACGGCAAAAGAAAATGGCGCGCTTCATTCCGCTGTTAGCCATACTGTTT 300
   1                          M  A  R  F  I  P  L  L  A  I  L  F   12

 301 ATGGCGTTTTGTAGCGGTGTGGCTCAAGCGTTACCAACAACTGATGAACGGGAGCAGAAC 360
  13 M  A  F  C  S  G  V  A  Q  A  L  P  T  T  D  E  R  E  Q  N   32

 361 GAAGTAGGACCGCCGGAAACTGGTACCATTTCCGACATGACGTCTGAGTGGGATCTGGCG 420
  33 E  V  G  P  P  E  T  G  T  I  S  D  M  T  S  E  W  D  L  A   52

 421 CTCCACAGGTCCAGGTCGACGGACTGTCGCCTGGCGGGCTATATAAAGAGAATCGCGATG 480
  53 L  H  R  S  R  S  T  D  C  R  L  A  G  Y  I  K  R  I  A  M   72

 481 CCATGGTGCCACACGGCTACCGTTCTCATCAACGCCTGCCGCGGACACTGCGAATCACAG 540
  73 P  W  C  H  T  A  T  V  L  I  N  A  C  R  G  H  C  E  S  Q   92

 541 ACCTCCCTGTCGTCGTACGCCACCGTGCAGGCCTCGGGTGGACAGCAGGTCTACACCACC 600
  93 T  S  L  S  S  Y  A  T  V  Q  A  S  G  G  Q  Q  V  Y  T  T  112

 601 AGGGGGAGCTGCTGCACCATAGCAACTACACATCAGGTGTCGTTCAACGTGGTGTGTTGG 660
 113 R  G  S  C  C  T  I  A  T  T  H  Q  V  S  F  N  V  V  C  W  132

 661 AACGGTGTCAGGACCTATACTATCGAGTCTGCTGCCAGCTGCGCATGCGGAGTATGTGAC 720
 133 N  G  V  R  T  Y  T  I  E  S  A  A  S  C  A  C  G  V  C  D  152

 721 TACAACCCGTGATGTAGGGGGTCGGTCAGGTGCACCAGAGAGACTCTTAGATCAGATACT 780
 153 Y  N  P  *                                                   155

 781 CACGATATAGTTACGTGTGCTACAGGGCAGTTATGGACAATTCTATAACTAGACGGCGTT 840
 841 GGACAACATCACAAATAACAGGATTACATTCCATTCTTTTTTGTAATAAACAAGATGCAG 900
 901 GTACAGCTGACTTTATCTTGACTATGTAAGGCATATTATGTCGTCCTTTATGGTATGTTC 960
 961 CAAACACTGTGTGGTGCGGAGAGCCCATGTATACACCACGTCCATACGTAGAAGCGCCAC 1020
1021 CTACATTCAAGCATAGATCTTTTCTTGCACGTATTTGAGGAAGTTCGGGATATCGGGGAC 1080
1081 ATAATGTTTTCCTTAACAATTTTTTTGCCTCTTATTCAAATGTTCATCCGATTATATAGT 1140
1141 GTAAGGGACTGTCAGAAAAGCCCGTCAGTCACTTAGAAAACATTTCGATCTTTCACCCCA 1200
1201 ACATCTGCTTGGAGACGTTAACTGATTTCGTATGAAAACTTTGGCCATTTGACTGTTCAC 1260
1261 CCGCTTAACTCTTTACTATTTCTATGTTTACTGCTGCTGAACAAAAACACGACAAAATAA 1320
1321 GATTCAAAAAAAAAAAAAAAAAAAAAA                                   1349
```

Figure 3-4. Nucleotide and amino acid sequences of ampGPA2LP. The underline indicates the signal peptide region.

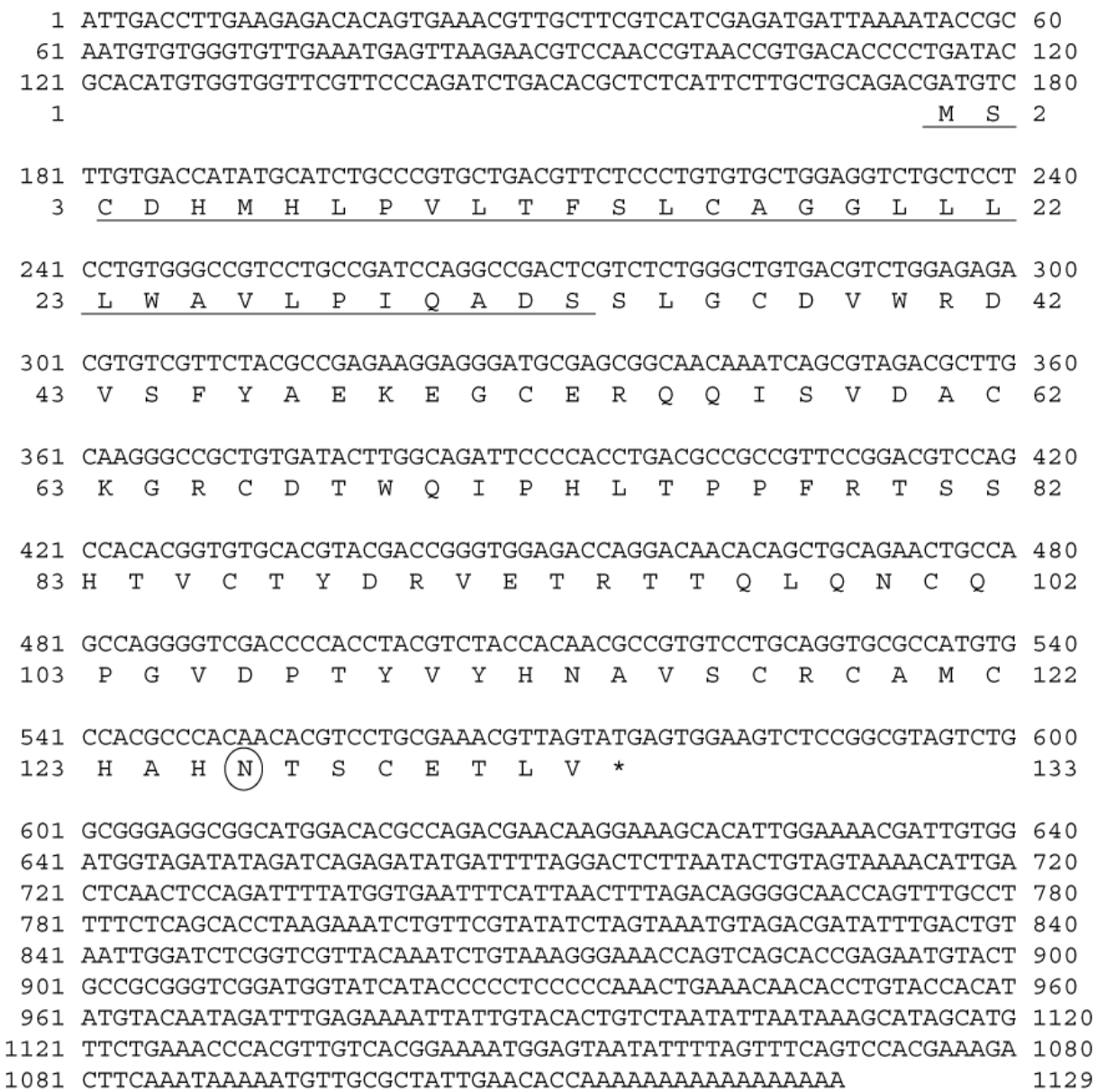

Figure 3-5. Nucleotide and amino acid sequences of ampGPB5. The underline indicates the signal peptide region. The potential N-linked glycosylation site is circled.

Fig. 3-6.

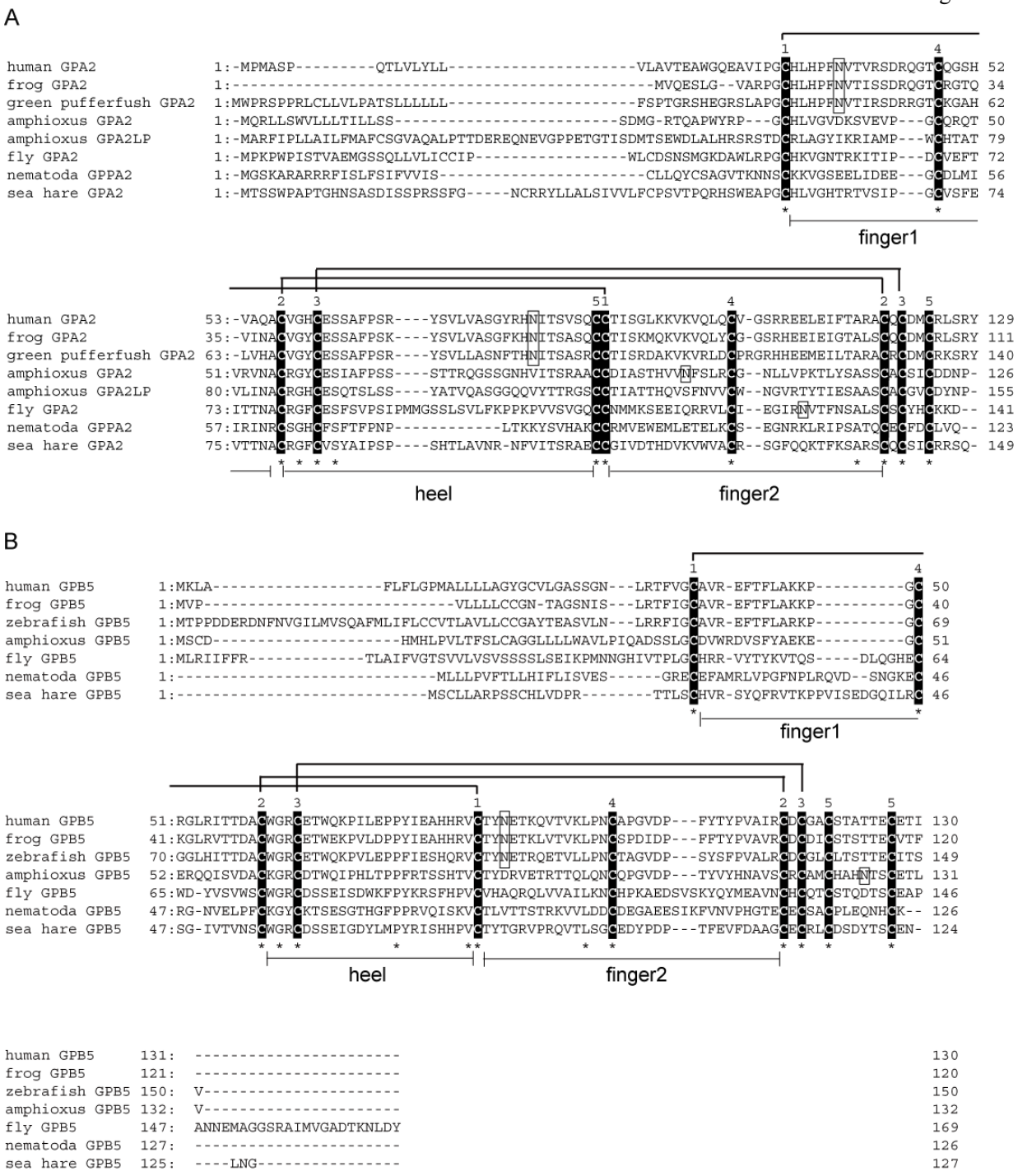

Figure 3-6. Alignments of amino acid sequences of ampGPA2, ampGPA2LP and ampGPB5. The amino acid sequences of ampGPA2, ampGPA2LP (A) and ampGPB5 (B) were compared with those of various animals. Identical amino acid residues are marked with asterisks and conserved cysteine residues are reversed. Numbers above Cys indicate the pairs of Cys, which are thought to form disulfide bonds. These pairs numbered as 1 to 3 which are connected with lines are necessary for constructing the cystine knot structure. Putative N-linked glycosylation sites are surrounded by squares. Fingers 1 and 2 indicate the outer parts of GPA2 and GPB5 molecules that forms β-strands, and the heel indicates the outer part of α-helical structure (see Figure 3-1).

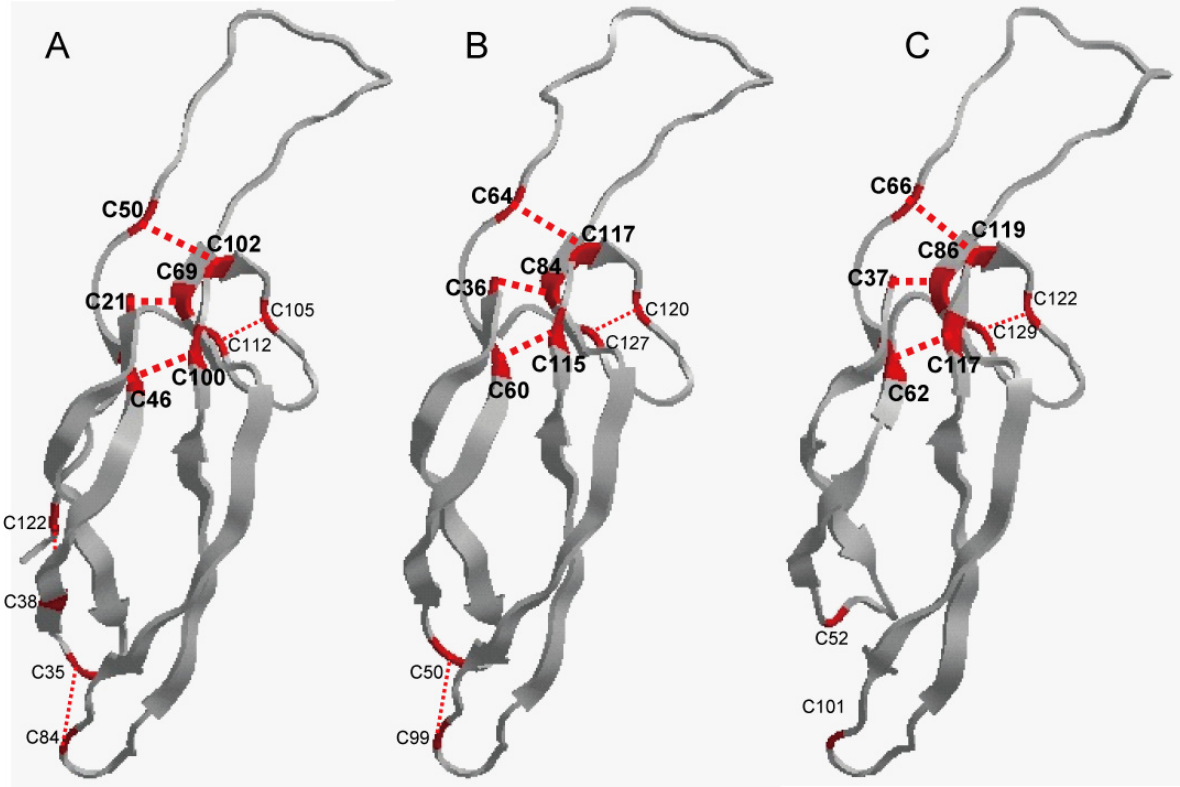

Figure 3-7. Comparison of the structures of human GPB5, ampGPB5 and human FSHβ subunits. (A) Structural model of human FSHβ. Accession code of the model is 1XWD in the Protein Data Bank. (B) Structural model of human GPB5 constructed by a protein modeling with the structure of human FSHβ. (C) Structural model of ampGPB5 constructed by a protein modeling with the structure of human FSHβ. Red-colored positions of these models indicate cysteine residues forming disulfide bonds, which are indicated by dotted red lines. The cysteine residues labeled with bold letters are necessary to form disulfide bonds and construct the cystine knot structure.

Fig. 3-8.

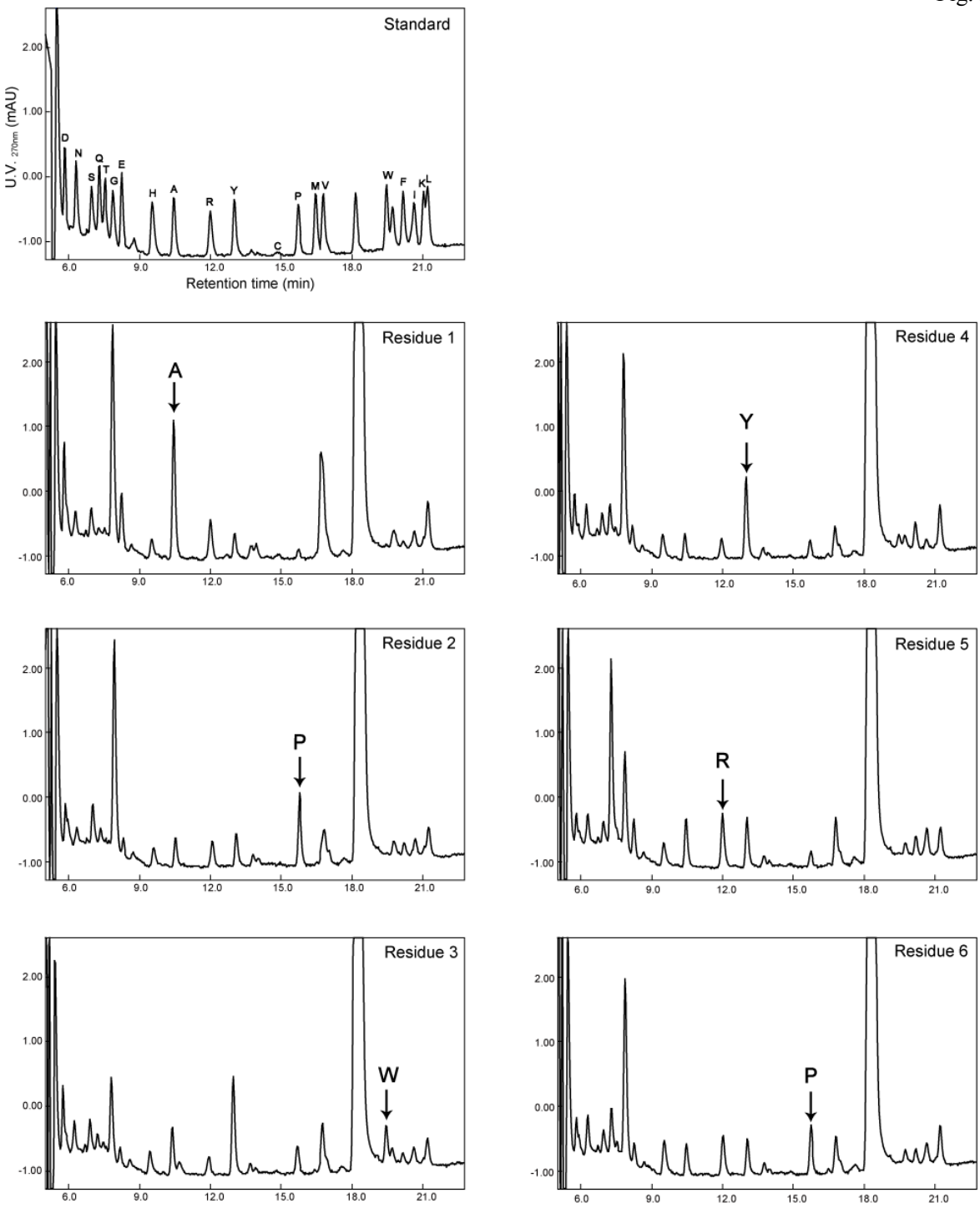

Figure 3-8. Sequence analysis of ampGPA2 by Edman degradation. The top panel shows PTH standard (Applied Biosystems) for Edman cycle 1. Cysteine residue was difficult to be detected in this setting. Residues 1 to 6 indicate amino acid profiles obtained by six Edman cycles following the first cycle for the PTH standard. Confirmed amino acids are arrowed in each panel. The amino acid residues are represented as the one-letter amino acid abbreviation. The first amino acid residue in the sequence of ampGPA2 was confirmed to be Ala25 (see Figure 3-3).

Fig. 3-9.

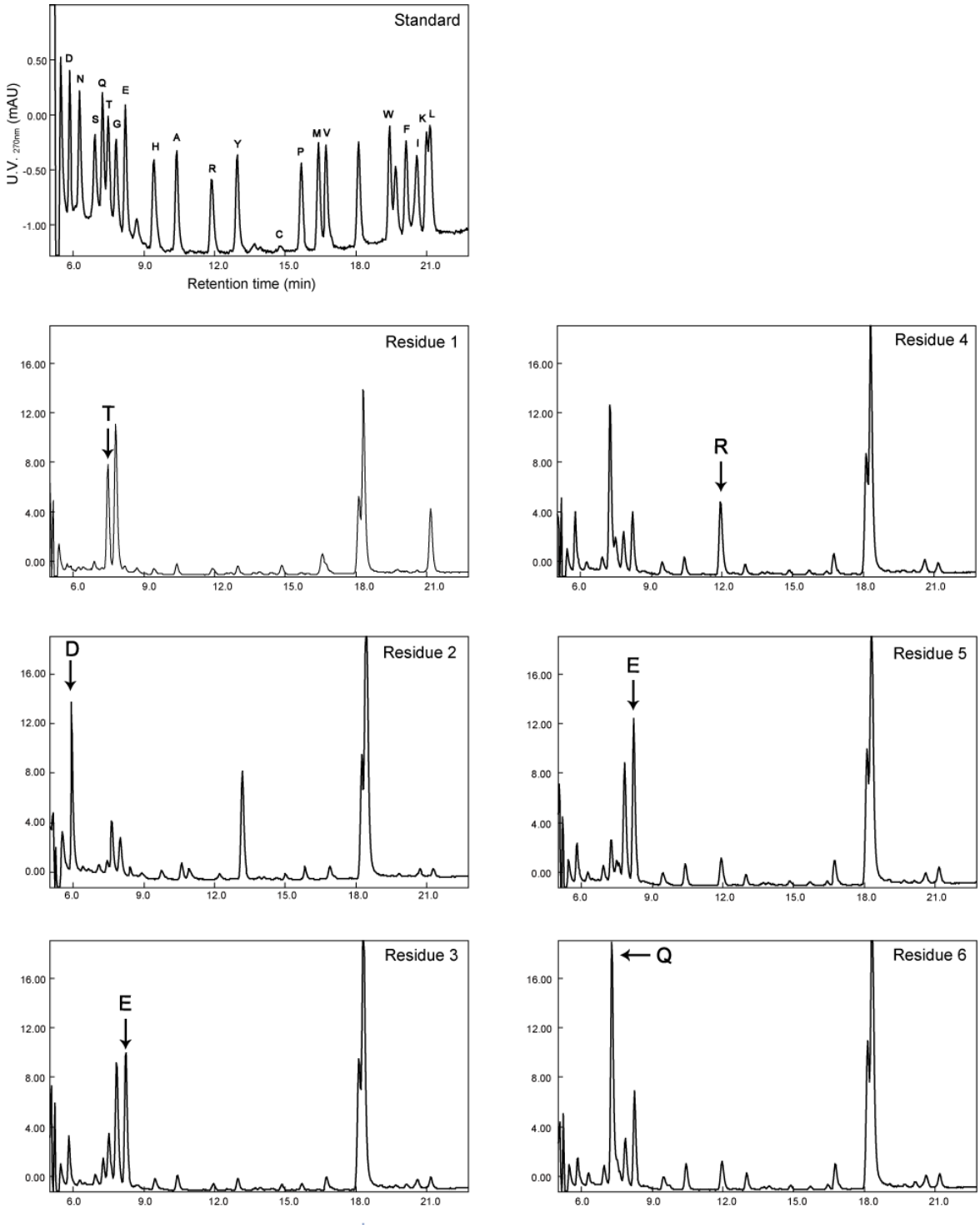

Figure 3-9. Sequence analysis of ampGPA2LP by Edman degradation. The top panel shows PTH standard (Applied Biosystems) for Edman cycle 1. Cysteine residue was difficult to be detected in this setting. Residues 1 to 6 indicate amino acid profiles obtained by six Edman cycles following the first cycle for the PTH standard. Confirmed amino acids are arrowed in each panel. The amino acid residue are represented as the one-letter amino acid abbreviation. The first amino acid residue in the sequence of ampGPA2LP was confirmed to be Thr26 (see Figure 3-4).

Fig. 3-10.

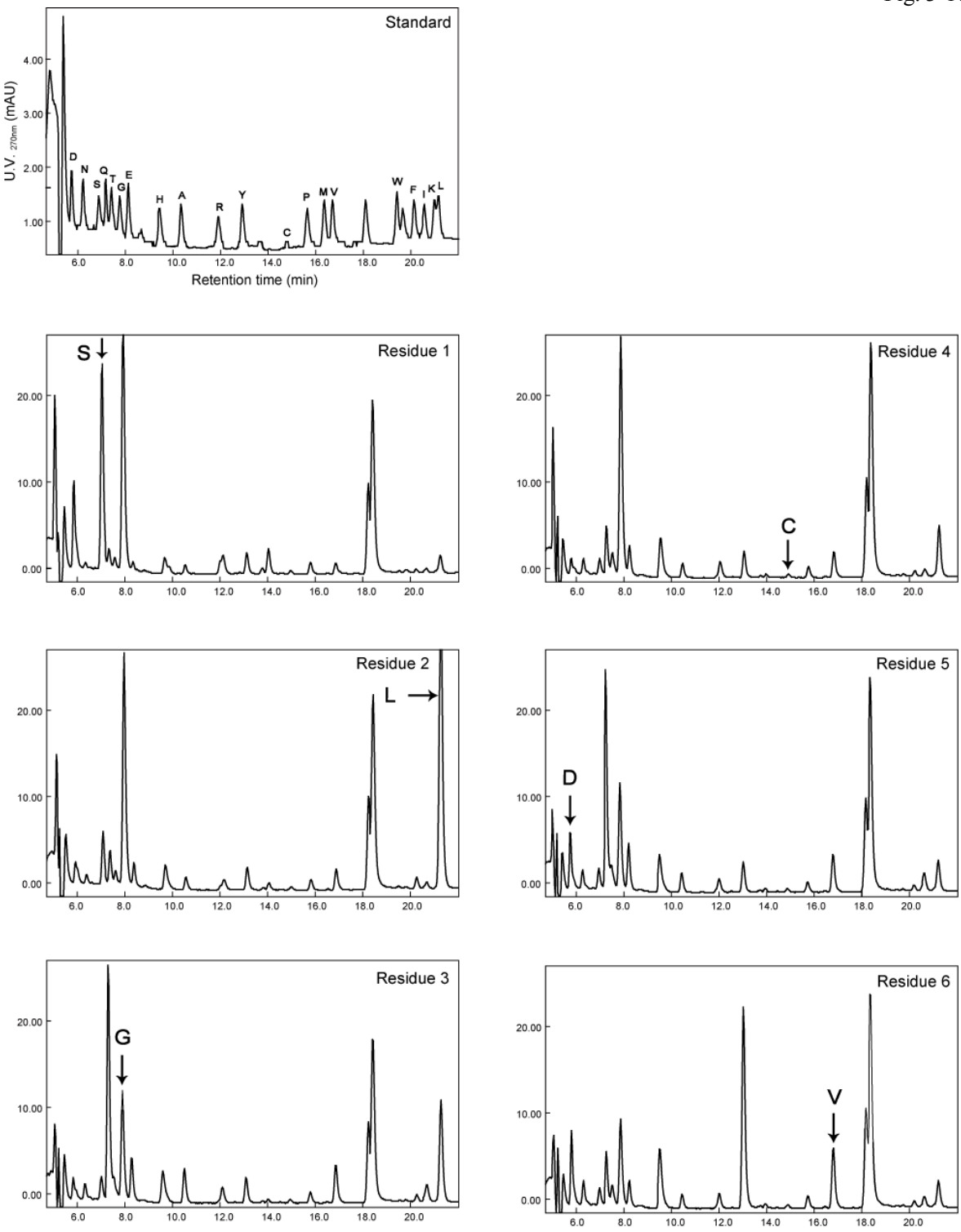

Figure 3-10. Sequence analysis of ampGPB5 by Edman degradation. The top panel shows PTH standard (Applied Biosystems) for Edman cycle 1. Cysteine residue was difficult to be detected in this setting. Cysteine in the panel of Residue 4 shows small peak. Residues 1 to 6 indicate amino acid profiles obtained by six Edman cycles following the first cycle for the PTH standard. Confirmed amino acids are arrowed in each panel. The amino acid residues are represented as the one-letter amino acid abbreviation. The first amino acid residue in the sequence of ampGPB5 was confirmed to be Ser34 (see Figure 3-5).

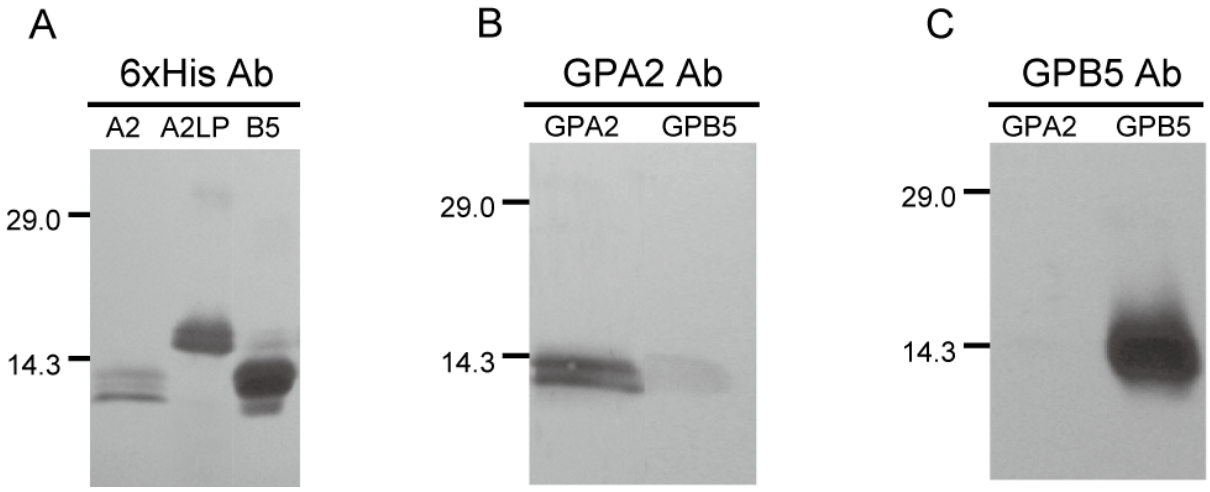

Figure 3-11. Western blot of recombinant ampGPA2, ampGPA2LP, and ampGPB5. (A) Recombinant ampGPA2, ampGPA2LP and ampGPB5 attached with His-Tag were purified and detected with anti-6x His antibody. (B, C) AmpGPA2 and ampGPB5 subunits were specifically detected with anti-GPA2 and anti-GPB5 antisera, respectively. Ab, antiserum.

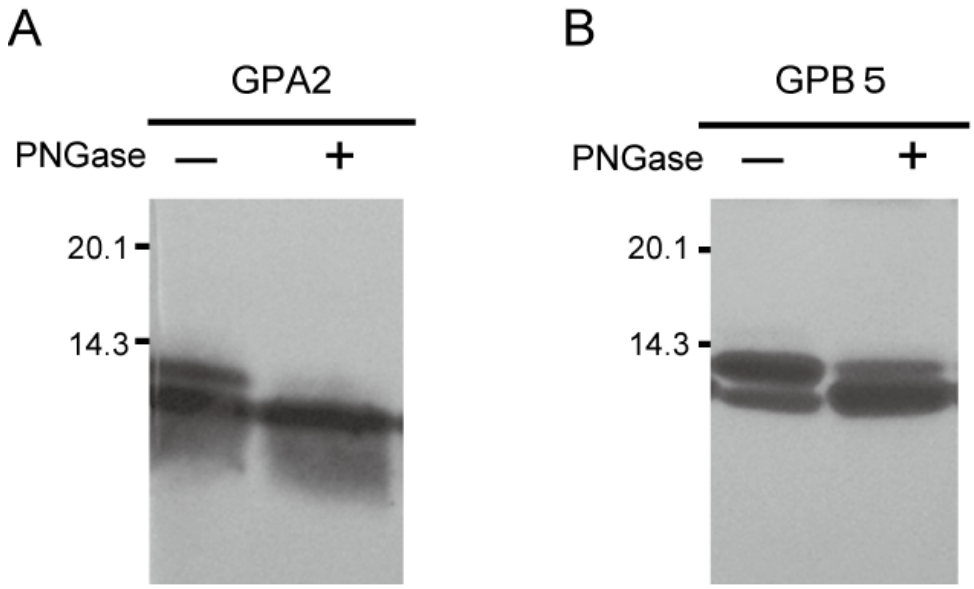

Figure 3-12. Glycosydase treatment of ampGPA2 and ampGPB5. Recombinant ampGPA2 (A) and ampGPB5 (B) were incubated with or without PNGaseF. Reaction products were separated by SDS-PAGE, and the removal of N-linked oligosaccharides from the proteins were detected by the reduction of molecular size using anti- 6x His antibody.

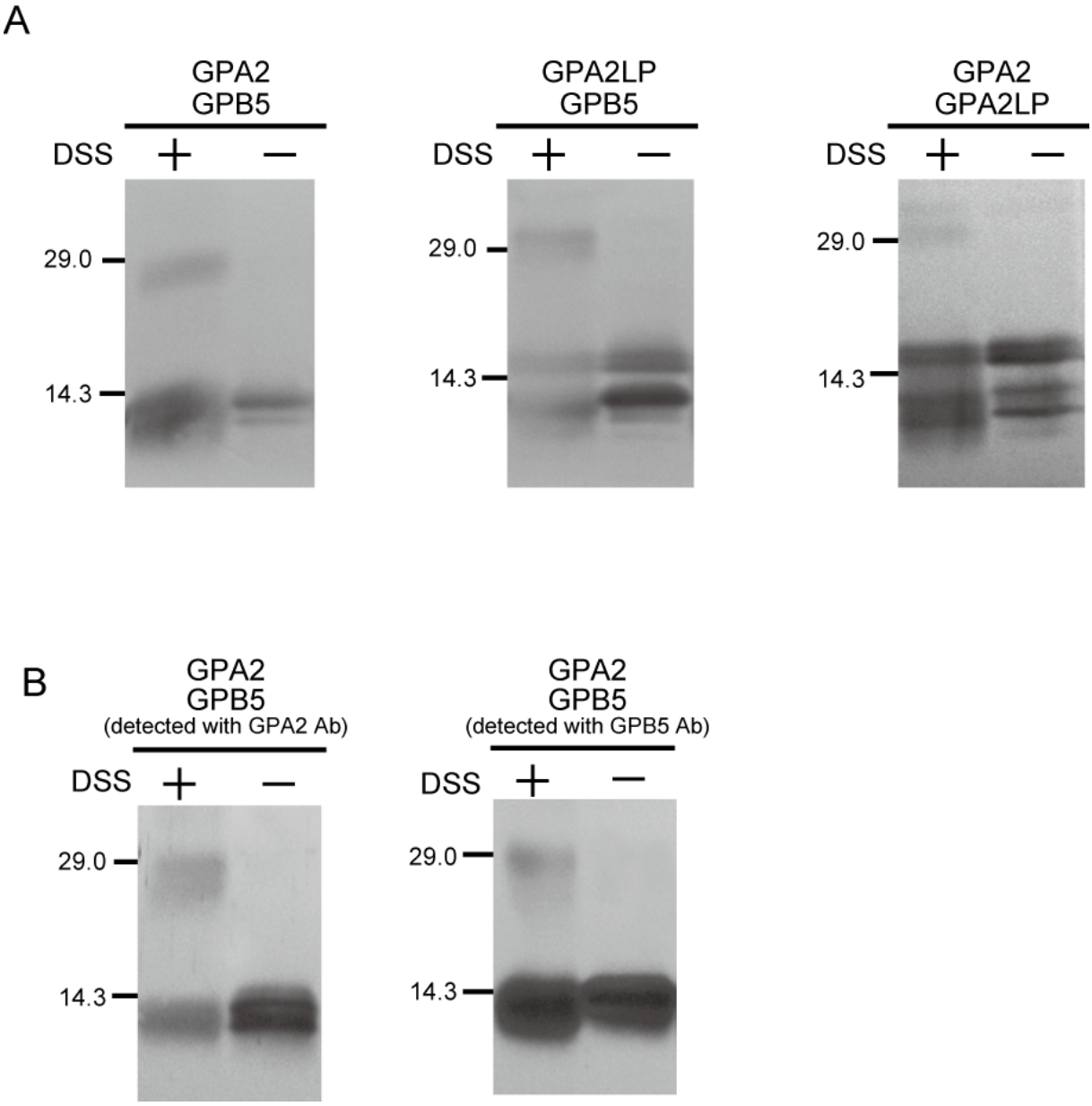

Figure 3-13. Dimer formation of ampGPA2, ampGPA2LP, and ampGPB5. (A) Mixtures used in a dimerization reaction were follows; ampGPA2 and ampGPB5 (left panel), ampGPA2LP and ampGPB5 (central panel), and ampGPA2 and ampGPA2LP (right panel). Mixtures were incubated with or without 1mM DSS and dimer formation was detected by Western blotting using anti-6x His antibody. (B) Mixtures of ampGPA2 and ampGPB5 were also detected by specific antisera against ampGPA2 (left panel) and ampGPB5 (right panel). Ab, antiserum.

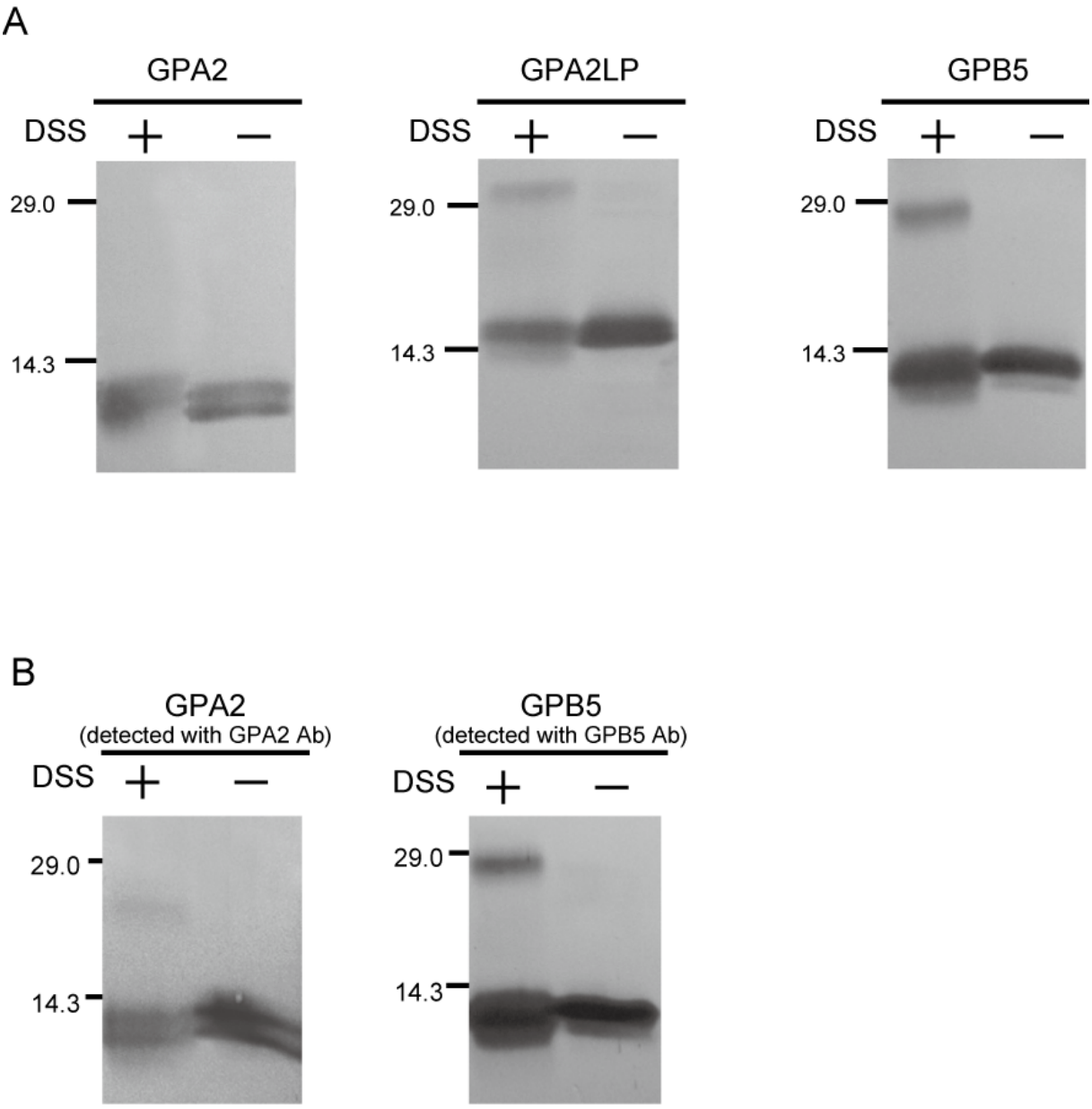

Figure 3-14. Dimer formation of ampGPA2, ampGPA2LP, and ampGPB5. (A) Each of ampGPA2 (left panel), ampGPA2LP (central panel) and ampGPB5 (right panel) was incubated with or without 1mM DSS. Reaction products were separated by SDS-PAGE, and detected by Western blotting using anti-6x His antibody. (B) Reaction products of ampGPA2 (left panel) and ampGPB5 (right panel) were detected by specific antisera against ampGPA2 and ampGPB5, respectively. Ab, antiserum.

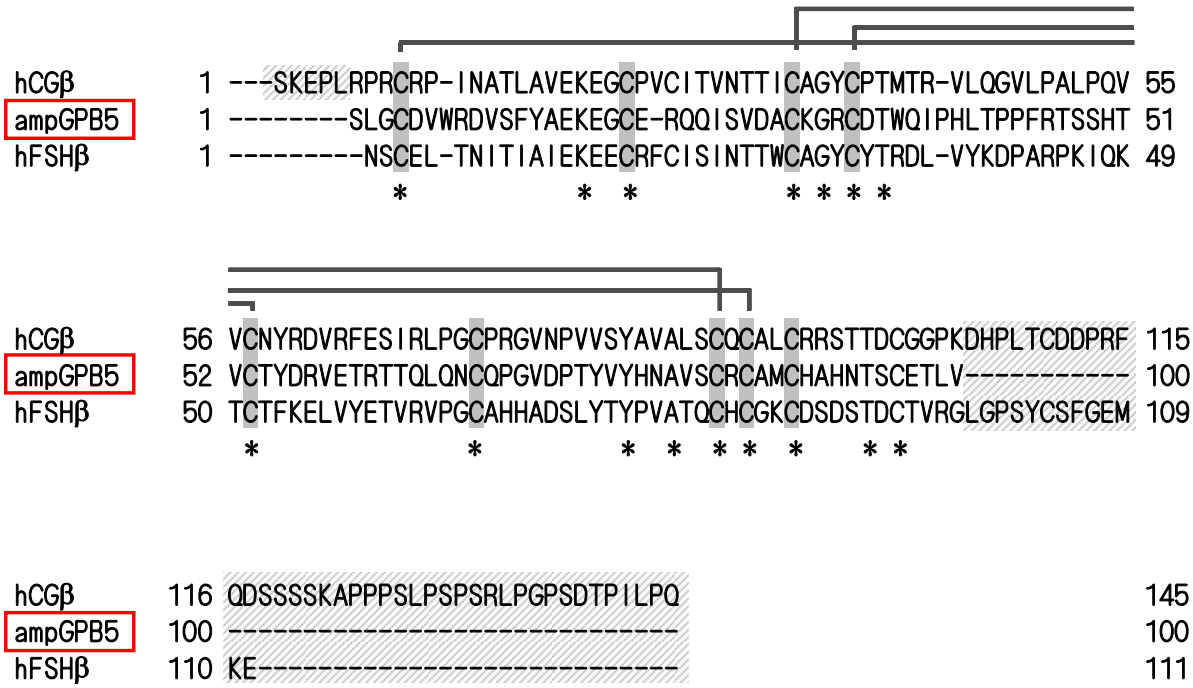

Figure 3-15. Alignment of human CGβ and FSHβ with amphioxus GPB5. The conserved cysteine residues are shaded. Portions of amino acid residues in hCGβ and hFSHβ longer than ampGPB5 are shaded with hatched lines. The disulfide bonds necessary for the cystine knot structure are indicated by lines which connect two cysteine residues above the alignment. amp, amphioxus; h, human.

Chapter 4

The Structures and Phylogenetic Relation of Amphioxus GPH Subunit Genes

4-1 Introduction

FSH, LH and TSH, members of the pituitary GPH family, are heterodimers of common α subunit and specific β subunit. Many researchers were involved in the investigation of the molecular evolution of vertebrate GPH subunits, in particular the divergence of genes for β subunits of FSH, LH, and TSH (e.g., Li et al., 1998; Querat et al., 2000). Sower et al. (2006) recently identified one GTHβ gene in lamprey, confirmed its expression in the pituitary, and then proposed that an ancestral glycoprotein hormone gave rise to only one GTH in ancestral lamprey, although this idea was soon turned over by the identification of cDNAs encoding α and β subunits of GTH in hagfish (Nozaki et al., 2006). Nevertheless, it seems to be true that, during the early evolution of gnathostomes before the radiation of vertebrates, an ancestral gene for the GPH family gave rise to genes for LH, FSH, and TSH evidently through duplications of the ancestral gene (Fig. 4-1) (Querat et al., 2001; Sower et al., 2006, 2009). However, evident orthologs of the genes for α and β subunits of pituitary GPH were not identified yet in the genomes of close extant relatives of the vertebrates (Blair and Hedges, 2005; Delsuc et al., 2006; Dunn et al., 2008) and any other invertebrates.

As mentioned in Chapter 3, similar nucleotide sequences of recently discovered GPH subunit genes, *GPA2* and *GPB5*, are present not only in vertebrates but also in invertebrates. For example, cDNAs encoding GPA2 and GPB5 were cloned from fruit fly (Sudo et al., 2005). Sequences similar to *GPA2* and *GPB5* genes were found in the genome databases of nematode (Park et al., 2005) and several protostomes (Dos Santos et al., 2009). On the basis of their global structural similarity with putative GPA and GPB subunits, they are considered to be phylogenetically related to pituitary GPH subunits. However, any molecular phylogenetic studies did not confirm their paralogous or orthologous relationships.

A survey of the amphioxus genome showed that they have only one set of homologous genes for pituitary GPH subunits, GPA2 and GPB5. This result indicates that genes

encoding amphioxus GPA2 and GPB5 are appropriate targets for elucidating the evolution of vertebrate GPH subunits. The phylogenetic position of amphioxus as an adjacent ancestor of chordates (Putnum et al., 2008) also enhances usefulness of genomic information of amphioxus in an evolutional study. In this chapter, I therefore carried out molecular phylogenetic analyses of GPH subunit genes to clarify relations of thyrostimulin subunit genes with pituitary GPH subunit genes. For this aim, I compared first the structures of amphioxus and human genes for GPA2 and GPB5 subunits and also those of human genes for pituitary GPH subunits. Subsequently, I analyzed conserved synteny of these genes, and compared them among amphioxus, human, and other vertebrates.

4-2 Materials and Methods

4-2-1 Comparison of the structures of GPA2 and GPB5 genes

Genomic DNA of *B. belcheri* was extracted from a whole body of single amphioxus with QuickGene-800 (Fujifilm, Kyoto, Japan) according to the manufacturer's instructions. Major portions of *GPA2* and *GPB5* genes were then amplified from the genomic DNA by PCR with a set of specific primers which were used for amplification of full-length cDNAs encoding GPA2 and GPB5 (see Chapter 3, Table 3-1). Nucleotide sequences of PCR products were determined with an ABI 3130 Genetic Analyzer (Applied Biosystems), and were compared with the sequences of human *GPA* and *GPB* genes that were downloaded from the Entrez Gene in the NCBI database (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene).

4-2-2 Phylogenetic analysis

Deduced amino acid sequences of amphioxus GPA2, GPA2LP and GPB5 (see Chapter 3) were aligned with those of GPAs and GPBs in both invertebrates and vertebrates by use of

the Clustal W program (Thompson et al., 1994). An unrooted phylogenetic tree of GPHs was then constructed by the neighbor-joining method along with calculation of the evolutionary distances by Poisson correction with the MEGA version 3.1 software (Kumar et al., 2004).

4-2-3 Analysis of the conserved synteny

Analyses of the conservedness of gene orders in the vicinities of GPH genes were carried out as described by Larsson et al. (2008) with several modifications. Briefly, the gene orders were investigated by use of *B. floridae* draft genome assembly version 1.0 in the JGI and the Ensembl human genome database v.52.36n. A Blastx search was applied to identify the genes annotated on the *B. floridae* scaffold 8, which contains *GPA2* and *GPB5* genes, and also location of corresponding genes in the human genome. Among the best hits on the human chromosomes, TSHβ is located on the chromosome 1, GPA1 on the 6, FSHβ on the 11, GPB5 on the 14, and LHβ on the 19, and syntenies of their neighboring genes were analyzed. Subsequently, the corresponding genes encoding the family proteins which locate on the vicinities of GPH genes of the *B. floridae* genome were verified in human chromosomes. In addition, these correspondent genes encoding family proteins of mouse, chick, green puffer fish and zebrafish were described from the Ensembl database information.

4-3 Results

4-3-1 Structures of amphioxus GPA2 and GPB5 genes and human GPH subunit genes

The structure of amphioxus *GPA2* gene, as well as *GPB5* gene, was elucidated by PCR amplification of genomic DNA and determination of sequence, followed by a comparison with the sequence of cDNA (see Chapter 3) to elucidate the exon-intron sturcture.

Amphioxus *GPA2* is comprised of four exons and three introns (Fig. 4-2A), while *GPB5* consists of two exons and one intron (Fig. 4-2B). Amphioxus *GPB5* gene is apparently smaller than human *GPB5* which is approximately 5 kbp. The 4 exon-3 intron structure of *GPA* genes and the 2 exon-1 intron structure of *GPB* genes are well conserved, whereas genes encoding β subunits of pituitary GPH have the 3 exon-2 intron structure, implicating that the number of exons in the hormone specific β subunits genes was increased after the ancestral GPH β subunit gene was diverged from *GPB* which is proposed by Sower et al. (2009).

4-3-2 Phylogenetic tree

The phylogenetic tree in Fig. 4-3 shows that each of amphioxus thyrostimulin subunits, amphioxus GPA2 (Fig. 4-3A, hereafter referred to as ampGPA2 as in the previous chapters) and GPB5 (Fig. 4-3B, referred to as ampGPB5), falls into the corresponding cluster. Similarly, the GPA2 cluster includes ampGPA2LP, which may have diverged from ampGPA2.

4-3-3 Analysis of the conserved synteny

My analyses of conserved gene orders in the vicinities of GPH subunit genes revealed an interesting fact that *ampGPA2* and *ampGPB5* genes are tandemly arranged on the scaffold 8 in the genome database of *B. floridae*, although *GPA2LP* gene is located on the scaffold 2. The order of genes near *ampGPB5* and *ampGPA2* genes on the scaffold 8 of *B. floridae* is comparable to those on human chromosomes (Fig. 4-4A); homologs of all or some of the seven *B. floridae* genes in the vicinity of *ampGPA2* and *ampGPB5* genes are found on human chromosomes 1, 11, 14, and 19, where GPH subunit genes exist. However, none of these seven genes were found on the scaffold 2, where *ampGPA2LP* was located. Homologs of the seven neighboring genes were not found on the human chromosome 6, where *GPA1* gene was located.

The conservedness of the order of genes for family proteins was further investigated in the genomes of mouse, chick, green pufferfish and zebrafish. The order of genes near the *GPB* gene is well conserved in the chromosomes where *GPB5* gene was located, particularly in the human chromosome 14 and the mouse chromosome 12 (Fig. 4-4B), although a few genes were deleted or moved onto other chromosomes in green pufferfish, zabrafish and amphioxus. The conservedness of syntenies in terms of genes for pituitary GPH subunits, i.e., TSHβ, both FSHβ and GPA2, and LHβ, was confirmed between human and mouse (fig4-5); however, the conservedness is rather low in green pufferfish and zebrafish, and some of the paralogous genes were located on separate chromosomes in green pufferfish and zebrafish.

4-4 Discussion

Molecular analyses in this chapter first showed that exon-intron structures of GPA genes are well conserved, whereas the number of intron increased in vertebrate pituitary GPB genes. The phylogenetic tree then showed that amphioxus GPA2 and GPB5 belong to corresponding clusters, but not to the clusters of vertebrate GPH subunits. Furthermore, the syntenies of *GPA2* and *GPB5* genes in amphioxus (*B. floridae*) are similar to those of GPH subunit genes in human.

4-4-1 Gene structure of amphioxus GPA2 and GPB5

The gene organization of *GPA2* and *GPB5*, in particular the exon-intron structure, is well conserved between amphioxus and human. Similarly, the gene organization of *GPA2* and *GPB5* is well conserved in distinct bilaterian phyla except for insects (Dos Santos et al., 2009). Since the exon-intron structure is known to be important for understanding of

molecular evolution of certain genes (Kotani et al., 1986), similarities of gene structures seen in my study probably indicate that a gene for α subunit of pituitary GPH of vertebrates and ampGPA2 originated from a common ancestral gene. Genes encoding β subunits of pituitary GPH and ampGBP5 gene also derived from another ancestral gene and increased the number of exons before diversion of ancestral GPH β subunit genes.

4-4-2 Molecular evolution of GPH subunits

Analyses of syntenies showed that the order of genes in the vicinities of *ampGPB5* and *ampGPA2* genes on the scaffold 8 of *B. floridae* is comparable to those near human genes encoding FSHβ, LHβ, and TSHβ. This fact supports the abovementioned idea that *ampGPA2* and *ampGPB5* genes existed in an ancestral cephalochordate are ancestral genes of human GPH subunit genes. Since Ohno (1970) hypothesized that major genome duplications occurred around the origin of vertebrates, three genes each for *FSHβ*, *LHβ*, and *TSHβ* would be derived from common ancestor with amphioxus *GPB5* gene by interchromosomal arrangement of genes that occurred after two duplication events in the vertebrate lineage.

Similar to *ampGPB5* and *ampGPA2* genes on the scaffold 8 of *B. floridae*, *GPA2* and *GPB5* genes were located nearby in the genomes of several non-vertebrate species (Dos Santos et al., 2009). It is therefore highly probable that not only β but α subunit gene for vertebrate GPH arose from the locus of tandem *ampGPB5* and *ampGPA2*. The noticeable is that the synteny of tandem *ampGPB5* and *ampGPA2* is comparable to that of tandem *FSHβ* and *GPA2* on human chromosome 11. Since a putative model of tertiary structure of GPB5 was completely generated by the use of the crystal structure of hFSHβ as a model, the synteny of *FSHβ* is considered to be highly conservable.

Sower et al. (2009) proposed that an ancestral glycoprotein hormone gave rise to

lamprey GTH and then to the glycoprotein hormone family that produced LH, FSH and TSH early in the evolution. If it is true, pituitary GPHs generated just after the emergence of agnathans with the acquisition of the pituitary gland as the unique endocrine organ in vertebrates.

4-5 Conclusion

My study in this chapter aimed to clarify whether subunits of amphioxus thyrostimulin are ancestral to those of vertebrate GPHs. The study on gene structures showed that the exon-intron structures of GPAs are well conserved, whereas the number of intron increased in vertebrate pituitary GPBs. The present phylogenetic tree showed that amphioxus GPA2 and GPB5 belong to corresponding clusters, but not to the vertebrate pituitary GPH subunit groups. Conserved syntenies of *GPA2* and *GPB5* genes in amphioxus (*B. floridae*) and GPH subunit genes in human demonstrated that amphioxus *GPA2* and *GPB5* genes and genes of human GPH subunit genes are originated from common ancestral genes. Because *GPA2* and *GPB5* genes in amphioxus and most of other invertebrates are located in the close vicinity, it can be hypothesized that vertebrate GPH subunit genes arose from the locus of amphioxus *GPA2* and *GPB5* by two rounds of whole genome duplications.

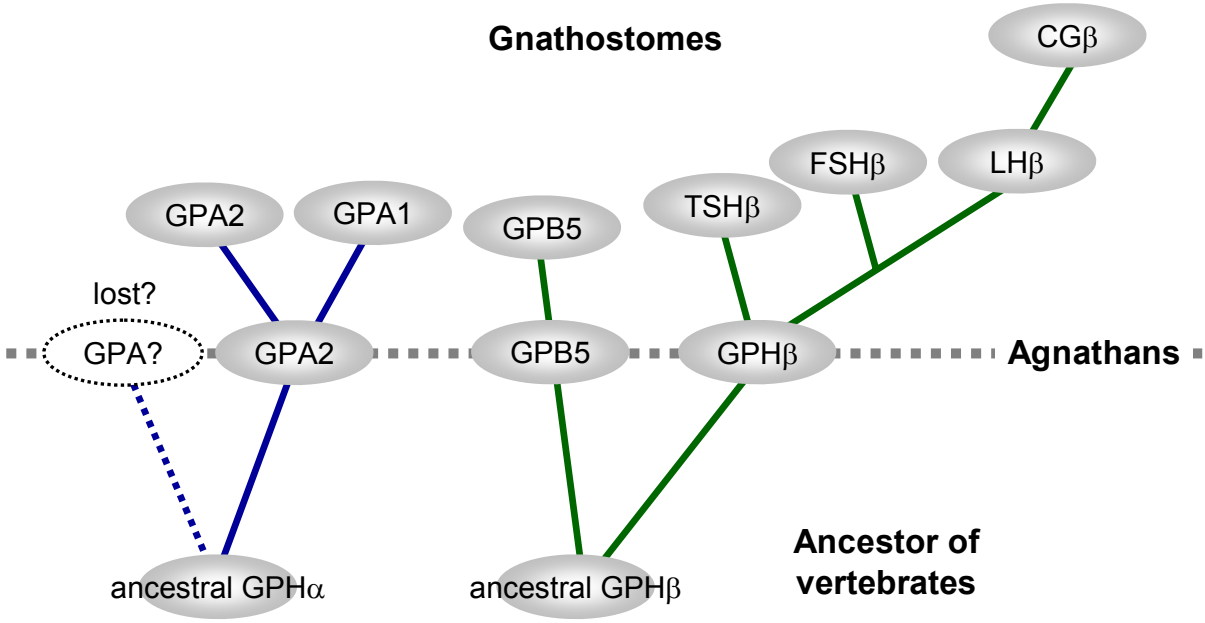

Figure 4-1. Scheme of the evolution of glycoprotein hormones (GPH) in vertebrates (modified from Sower et al., 2009). Agnathans including lamprey have one GPHα subunit (GPA2) and two GPHβ subunits (GPB5 and GPHβ). During the evolution from ancestral vertebrates to current vertebrates, an ancestral GPHβ was possibly diverged to GPB5 and GPHβ by the gene duplication. By the second gene duplication of GPHβ after the agnathan-gnathostome divergence, four GPHβ subunits (FSHβ, LHβ, TSHβ and CGβ) were evolved in gnathostomes. The evolution of ancestral GPHβ by gene duplication suggests that a gene for ancestral GPHα was also duplicated to form two α subunits in agnathans, although there is only one GPA2. The other GPHα might be lost in vertebrates.

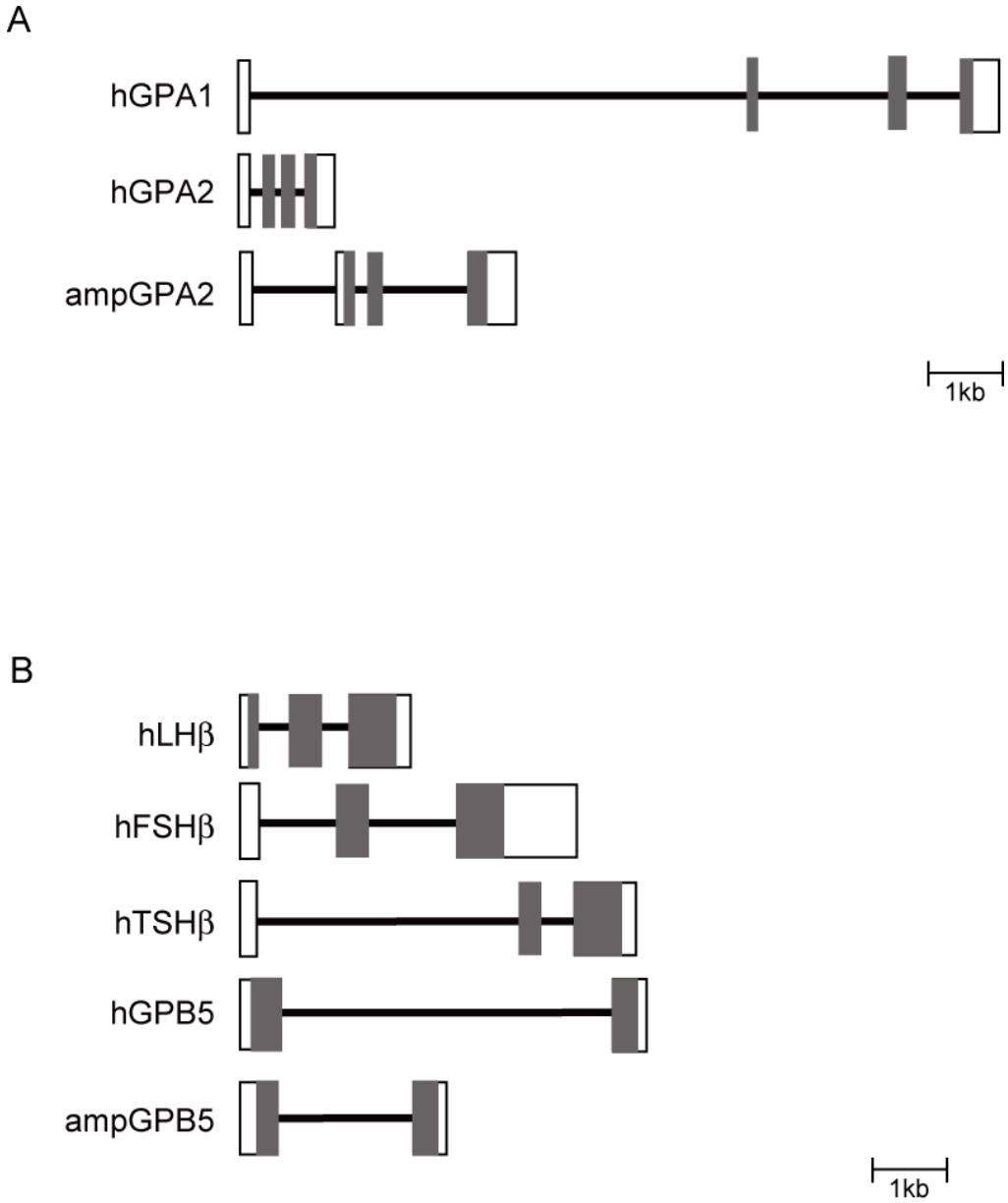

Figure 4-2. Comparison of structures of *ampGPA2* and *ampGPB5* genes with those of human GPHs. (A) Gene structure of α subunits. (B) Gene structure of β subunits. Squares indicate exons. Open squares indicate untranslated regions, while gray squares show translated regions. amp, amphioxus; h, human.

Fig. 4-3.

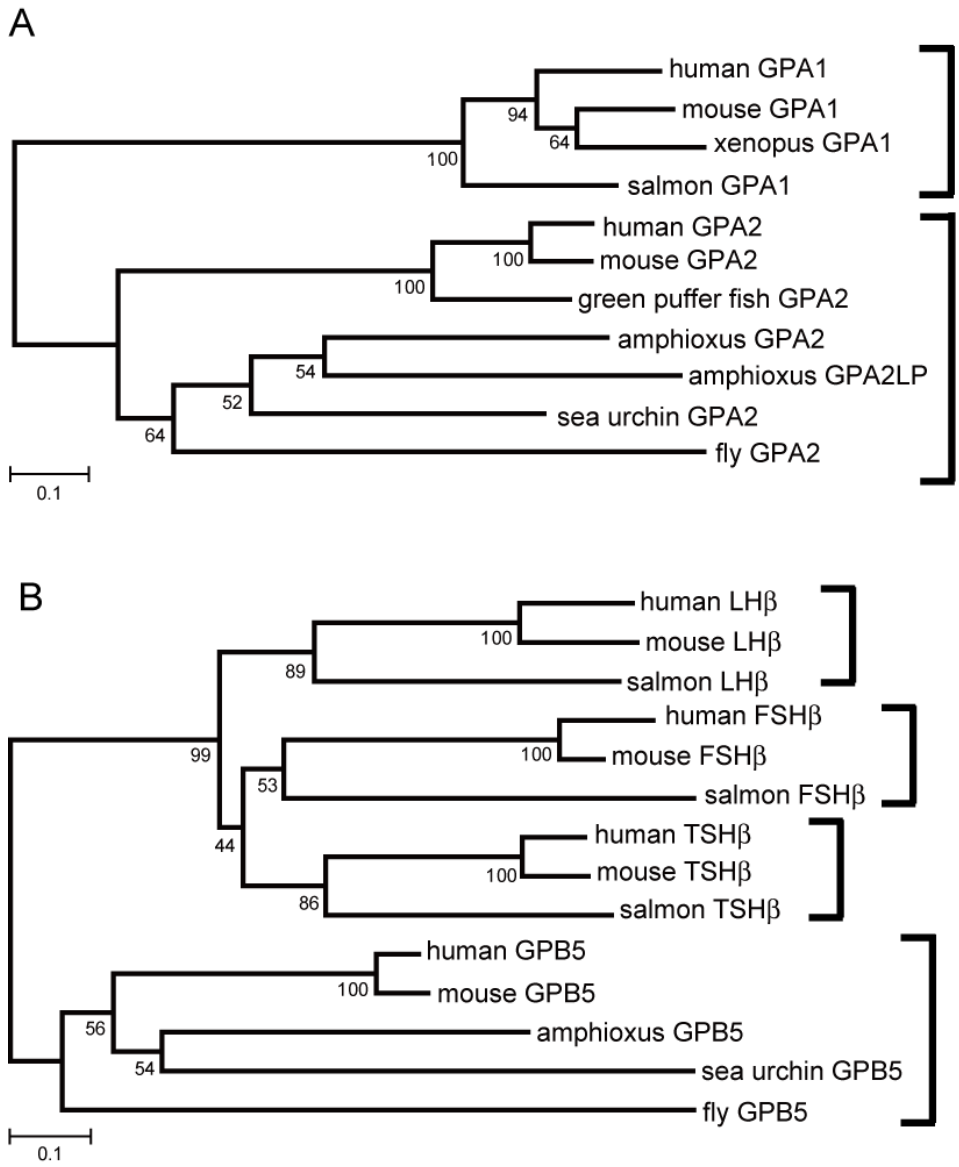

Figure 4-3. Unrooted molecular phylogenetic trees of GPH subunits constructed by the Neighbor-Joining method. (A) Unrooted tree of α subunits of GPHs. (B) Unrooted tree of β subunits of GPHs. The numbers beside the branches indicate bootstrap probabilities in the 1000 replication trials for construction of the tree. The bar represents an evolutionary distance (substitution/site) calculated by poisson correction. The DDBJ/EMBL/GenBank accession numbers of sequences used for analysis are as follows: human GPA1 (J00152); mouse GPA1 (J00643); xenopus GPA1 (L07619); salmon GPA1 (AB050834); human GPA2 (AF260739); mouse GPA2 (AF260740); green pufferfish GPA2 (Q4S0U2); fly GPA2 (AY940435); sea urchin GPA2 (15344); human LHβ (NM000894); mouse LHβ (NM008497); salmon LHβ (AB050836); human FSHβ (NM000510); mouse FSHβ (NP032071); salmon FSHβ (AB050835); human TSHβ (NM000549); mouse TSHβ (NM009432); salmon TSHβ (AF060566); human GPB5 (AF403430); mouse GPB5 (NM175644); fly GPB5 (AF403389); sea urchin GPB5 (DN791067).

Figure 4-4. Syntenic homologies in the vicinity of GPH subunit genes. (A) Synteny conservation of genes for GPHα and GPHβ subunits between amphioxus *Branchiostoma floridae* and human. Note that *ampGPB5* and *ampGPA2* locate in tandem on the amphioxus scaffold 8, as *FSHβ* and *GPA2* on the human chromosome 11. *AmpGPA2LP* is localized on the amphioxus scaffold 2, while human *GPA1* on the chromosome 6. Human *TSHβ* is on the chromosomes 1, *GPB5* on the 14, and *LHβ* on the 19. The names of the family protein registered in the gene model of *B. floridae* are shown on the top, and their IDs obtained from the *B. floridae* genome database in Joint Genome Institute are indicated in boxes on amphioxus scaffold 8. The numbers below boxes indicate the gene locus on amphioxus scaffold 8 and human chromosomes. chr, chromosome; Sc, scaffold. (B) Conserved syntenies of *GPB5* in the genomes of human, mouse, chick, green pufferfish, zebrafish, and amphioxus. The names of family proteins on the line of human chromosome 14 describe the gene models in human genome. The numbers in the boxes of vertebrate animals indicate the chromosome number in each animal, and that of amphioxus indicates the loci on scaffold 8. The numbers below boxes indicate the gene loci on the chromosome or scaffold. Un in chick and green pufferfish, unknown chromosome.

Fig. 4-4.

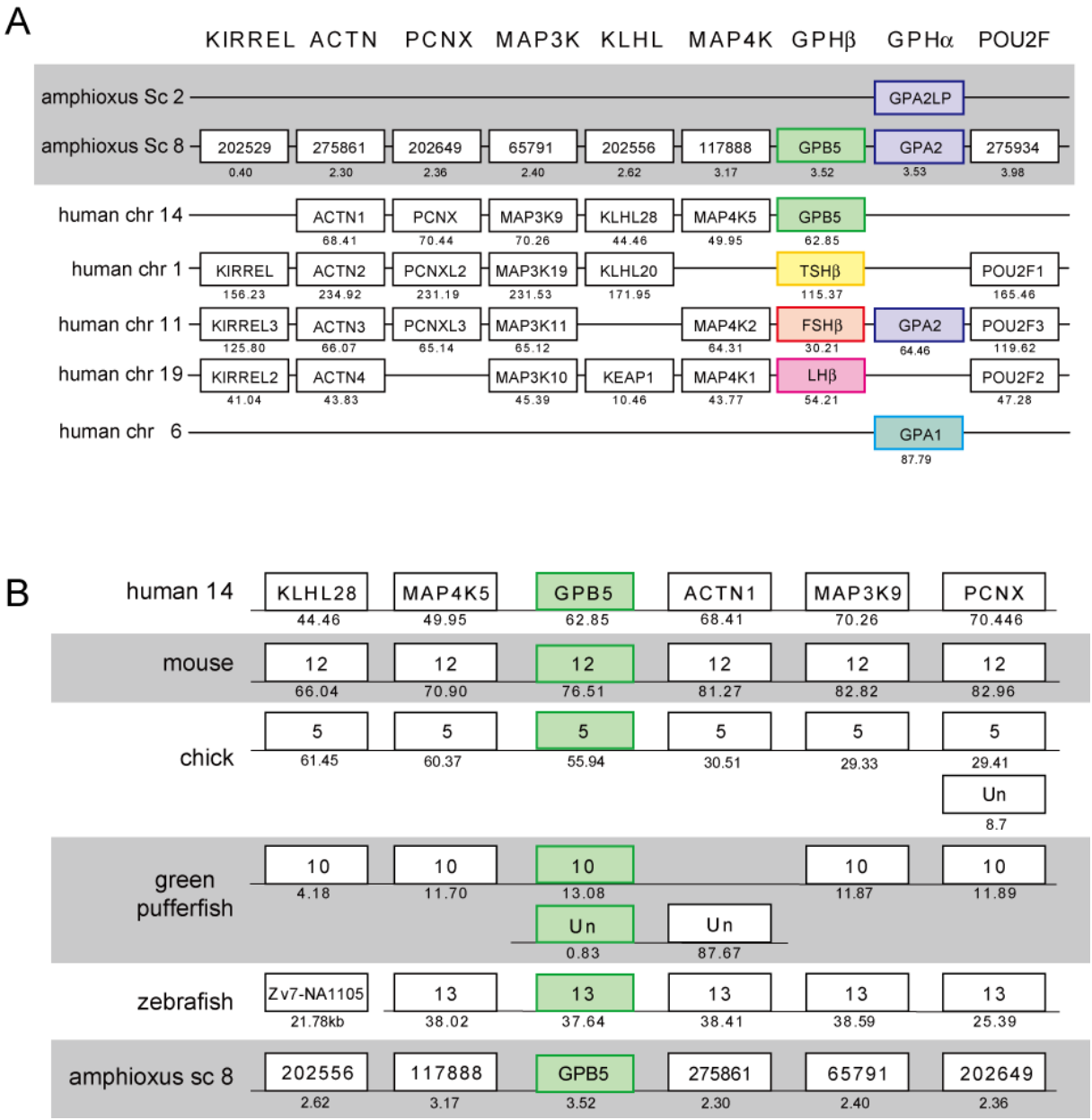
A
KIRREL ACTN PCNX MAP3K KLHL MAP4K GPHβ GPHα POU2F
amphioxus Sc 2
GPA2LP
amphioxus Sc 8
202529 275861 202649 65791 202556 117888 GPB5 GPA2 275934
0.40 2.30 2.36 2.40 2.62 3.17 3.52 3.53 3.98
human chr 14
ACTN1 PCNX MAP3K9 KLHL28 MAP4K5 GPB5
68.41 70.44 70.26 44.46 49.95 62.85
human chr 1
KIRREL ACTN2 PCNXL2 MAP3K19 KLHL20 TSHβ POU2F1
156.23 234.92 231.19 231.53 171.95 115.37 165.46
human chr 11
KIRREL3 ACTN3 PCNXL3 MAP3K11 MAP4K2 FSHβ GPA2 POU2F3
125.80 66.07 65.14 65.12 64.31 30.21 64.46 119.62
human chr 19
KIRREL2 ACTN4 MAP3K10 KEAP1 MAP4K1 LHβ POU2F2
41.04 43.83 45.39 10.46 43.77 54.21 47.28
human chr 6
GPA1
87.79
B
human 14
KLHL28 MAP4K5 GPB5 ACTN1 MAP3K9 PCNX
44.46 49.95 62.85 68.41 70.26 70.446
mouse
12 12 12 12 12 12
66.04 70.90 76.51 81.27 82.82 82.96
chick
5 5 5 5 5 5
61.45 60.37 55.94 30.51 29.33 29.41
Un
8.7
green pufferfish
10 10 10 10 10
4.18 11.70 13.08 11.87 11.89
Un Un
0.83 87.67
zebrafish
Zv7-NA1105 13 13 13 13 13
21.78kb 38.02 37.64 38.41 38.59 25.39
amphioxus sc 8
202556 117888 GPB5 275861 65791 202649
2.62 3.17 3.52 2.30 2.40 2.36

Fig. 4-5.

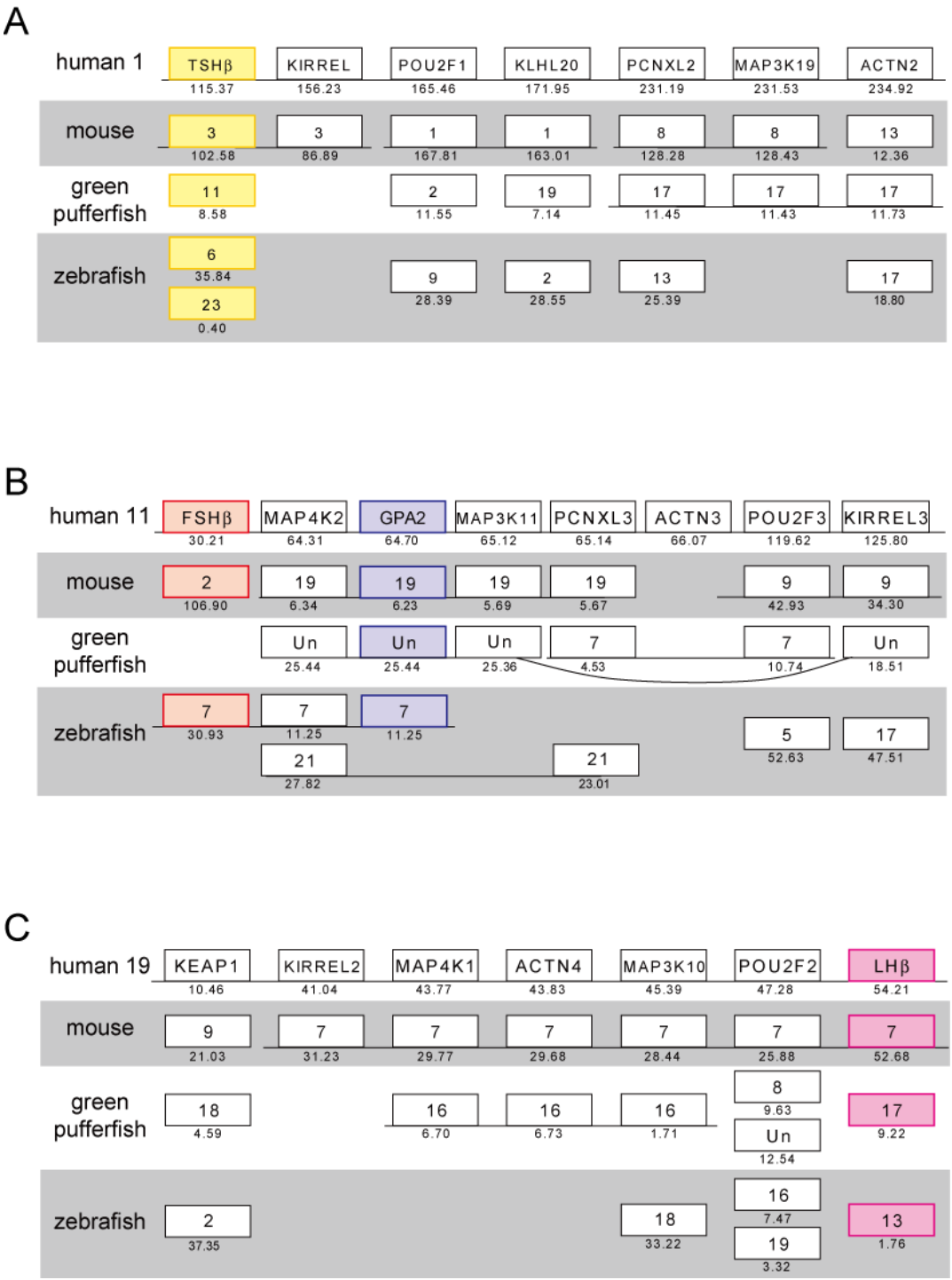

Figure 4-5. Syntenic homology of GPH subunit genes among various vertebrates, when compared with the gene orders on the human chromosome 1 that has *TSHβ* (A), the chromosome 11 that has *FSHβ* (B), and the chromosome 19 containing *LHβ* (C). Concerning the conserved synteny, see the text for explanation. The names of family proteins are indicated in the boxes of human chromosomes, while boxes of other animals indicate numbers of chromosomes. The orders of family proteins are arranged following those in human. The numbers below boxes indicate the locus of gene in a chromosome. Un in green pufferfish, unknown chromosome.

Chapter 5

Distribution of ampGPA2 and ampGPB5 in Amphioxus

5-1 Introduction

In the pituitary gland of tetrapod, TSH is sysnthesized in thyrotrophs, and FSH and LH are shynthesized mostly in the same gonadotrophs (Nakane, 1970; Moriarty, 1973, 1976), whereas immunohistochemical (IHC) and *in situ* hybridization (ISH) studies showed that FSH and LH are independently produced in different cells in the pituitary gland of teleost (Nozaki et al., 1990; Kagawa et al., 1998). These specific phenotypes of pituitary endocrine cells arise from a common primordium by sequential transcriptional regulation (Scully and Rosenfeld, 2002). It is generally accepted that GPHs are synthesized mainly in the pituitary gland, although several reports showed the presence of GPH subunits out of the pituitary gland (Croxatto et al., 1964; Emanuele et al., 1981; Hojvat et al., 1982; Hostetter et al., 1987).

Like pituitary GPHs, the presence of thyrostimulin subunits GPA2 and GPB5 was immunohistochemically demonstrated in pituitary cells (Nakabayashi et al., 2002; Okada et al., 2006). According to Okada et al. (2006), GPA2 and GPB5 are colocalized in corticotrophs in the pituitary gland of human; however, distribution of GPA2 and GPB5 are highly variable in other tissues than the pituitary gland. Expression of *GPA2* and *GPB5* genes are easily detectable by RT-PCR in diverse mammalian tissues, such as the brain, pancreas, placenta and gonads, as well as in the pituitary gland (Nakabayashi et al., 2002; Hsu et al., 2002; Nagasaki et al., 2006), although the levels of gene expression and distributions of two subunits are different. Human GPA2 mRNA is abundant in the pancreas, while GPB5 mRNA is detected mostly in the pituitary gland. In contrast, the amount of GPA2 is relatively high in the eye, while GPB5 is mostly expressed in the reproductive organs in rats (Nagasaki et al., 2006). These distribution patterns of thyrostimulin subunits are not consistent with the concept of pituitary GPHs.

The distribution of GPA2 and GPB5 are scarcely investigated in invertebrates. Sudo et al. (2005) reported that gene expression of fly *GPA2* and *GPB5* was detected by RT-PCR

through their life from an embryonic stage. Dos Santos et al. (2009) cloned cDNAs encoding *GPA2* and *GPB5* from *B. floridae*, and characterized expression of two subunit genes in embryonic and larval stages of amphioxus. The result showed that *GPA2* expression is restricted to particular areas, whereas *GPB5* is essentially ubiquitous through embryonic and larval stages. However, tissue distributions of two subunit genes are not determined. In this chapter, I investigated distributions of transcripts of amphioxus *GPA2*, *GPA2LP*, and *GPB5* in adult amphioxus using by RT-PCR and ISH.

5-2 Materials and Methods

5-2-1 RT-PCR for detection of tissue specific gene expression

The head, skin, gills including endostyle, muscle, testes, and ovaries of amphioxus were dissected out from mature adults with forceps under a dissecting microscope, immediately frozen in liquid nitrogen, and were stored at -80^{o}C until extraction of total RNA. The head was defined as the region from the tip of the anterior end to the velum. From each tissue, total RNA was extracted and reverse-transcribed by the methods previously described. PCR amplification of cDNA from dissected tissues was carried out under the following conditions.

GPA2 30 cycles (95^{o}C for 30 s, 65^{o}C for 30 s, 72^{o}C for 30 s), 72^{o}C for 3 min

GPA2LP 30 cycles (95^{o}C for 30 s, 56^{o}C for 30 s, 72^{o}C for 30 s), 72^{o}C for 3 min

GPB5 30 cycles (95^{o}C for 30 s, 60^{o}C for 30 s, 72^{o}C for 30 s), 72^{o}C for 3 min

Gene specific primers for *GPA2*, *GPA2LP*, and *GPB*5 are listed in Table 5-1. *EF-1*α, sequence of which was derived from first-strand cDNA synthesized from whole body total RNA of *B. belcheri*, was served to confirm intactness of cDNAs. The same amounts of PCR products (150 ng) were analyzed by electrophoresis in 2% agarose gels with ethidium bromide.

5-2-2 in situ hybridization (ISH)

Hybridization probes: To construct a molecular probe for *in situ* hybridization, cDNA encoding full length nucleotide sequence of *GPA2, GPA2LP, and GPB5* were used as the template, and amplified by PCR with a set of two primers:

for GPA2, GPA2-F: 5'-CGACCACCTTAAGCAATCAC-3' and

GPA2-R: 5'-TCCTGCATGAGGTTGTTGGA-3' (nt 8-1162; 1155 bp)

for GPA2LP, GPA2LP-F: 5'-CAATTTCCTTATCACGACAGAC-3' and

GPA2LP-R: 5'-GGATGAACATTTGAATAAGAGGC-3' (nt 97-1129; 1033 bp) and

for GPB5, GPB5-F: 5'-GTCCAACCGTAACCGTGACA-3' and

GPB5-R: 5'-TGGTGTTCAATAGCGCAACA-3' (nt 91-1113; 1023 bp).

PCR amplification of cDNA from dissected tissues was carried out under the following conditions.

GPA2 35 cycles (95^{o}C for 30 s, 65^{o}C for 30 s, 72^{o}C for 1 min), 72^{o}C for 4 min

GPA2LP 35 cycles (95^{o}C for 30 s, 58^{o}C for 30 s, 72^{o}C for 1 min), 72^{o}C for 4 min

GPB5 35 cycles (95^{o}C for 30 s, 65^{o}C for 30 s, 72^{o}C for 1 min), 72^{o}C for 4 min

The PCR products were purified with a QIAX II gel extraction kit (QIAGEN, Hilden, Germany) and subcloned into pCR4 TOPO plasmid vector (Invitrogen). Ten micrograms of purified recombinant plasmid DNAs were digested with Not I and Spe I (Takara), consecutively for the synthesis of sense and antisense RNA probes. Digoxigenin (DIG)-labeled sense and antisense RNA probes specific to each mRNA were generated with a DIG RNA labeling kit (Roche, Penzberg, Germany) according to the manufacturer's instructions. Labeled RNA probes were fragmented in fragmentation buffer (42 mM NaHCO3, 63 mM Na2CO3, 62.5 mM DTT) at 60^{o}C for 10 min, purified with Ultrafree-MC Centrifugal Filter Units (Millipore, Tokyo, Japan), and dissolved in 40 μl hybridization buffer.

Animals: Mature animals were fixed in 4% paraformaldehyde (PFA) in 0.1 M MOPS buffer

(pH 7.5) containing 0.5 M NaCl at 4^{o}C for 12 hrs. The specimens were dehydrated in a graded series of methanol/phosphate buffered saline, pH 7.4 (PBS); 25%, 50%, 75% and 100% for 30 min at room temperature, and stored at -30^{o}C until use. The nerve cords and heads that were used for a whole mount ISH (abbreviated as WISH) were dissected out from fixed heads in 80% methanol/PBS.

Whole-mount ISH: WISH was carried out as described by Ogasawara et al (2006) with minor modifications. The specimens in 80% methanol were rehydrated, and then washed with PBT. After a treatment with 20 μg/ml of proteinase K (Sigma) for 20 min, they were post-fixed with 4% PFA/PBT for 20 min. The specimens were placed in the InSitu tip (ALOKA, Tokyo, Japan), and were incubated in hybridization buffer (5x standard saline citrate (SSC), 50% formamide, 1% SDS, 50 μg/ml yeast tRNA, 50 μg/ml heparin, 0.1% CHAPS, 5mM EDTA) without probe at 50^{o}C for 1 h. Hybridization was carried out in a hybridization buffer, which contains DIG-labeled RNA probe (0.5-1.0 ng/μl), at 50^{o}C for 16 h. After hybridization, a series of washes were performed at 50^{o}C with wash buffer 1 (4x SSC, 50% formamide, 0.1% Tween 20), wash buffer 2 (2x SSC, 50% formamide, 0.1% Tween 20), wash buffer 3 (1x SSC, 50% formamide, 0.1% Tween 20), and finally wash buffer 4 (0.1x SSC, 50% formamide, 0.1% Tween 20) for 20 min. Afterward, the specimens were treated with 20 μg/ml RNaseA in PBST at 37^{o}C for 20 min and washed with a wash buffer 5 (1x SSC, 0.1% Tween 20). After color development for alkaline phosphatase with nitro blue tetrazolium/5-bromo-4-chloro- 3-indolyl phosphate (NBT/BCIP) and microphotography as whole mount specimens, the nerve cords were embedded in 5% agar and cut at 50 μm with a linear slicer (Douhan EM, Osaka, Japan).

5-2-3 Double color ISH

For double ISH staining, the anterior part of animal was fixed as described above,

embedded in OCT compound (Ted Pella), and cross-sectioned at 20 μm on a cryostat (Leica). A double color ISH was employed to clarify whether the same cells express both *GPA2* and *GPB5* genes. Thus, RNA probe for *GPB5* was labeled with DIG, and that for *GPA2* with fluorescein RNA labeling kit (Roche) according to the manufacturer's instructions. Labeled RNA probes were fragmented in fragmentation buffer as described above. The staining procedure followed that in Watakabe et al. (2007). In brief, frozen sections were treated with 20 μg/ml of proteinase K for 20 min at room temperature, and then incubated in a hybridization buffer containing 0.5-1.0 ng/μl DIG- and fluorescein- labeled riboprobes. Then, sections were washed with a series of wash buffers as same as WISH, and treated with 20μg/ml of RNase (Sigma) at room temperature for 20 min. Afterward, they were incubated with anti-fluorescein antibody conjugated with horseradish peroxidase (Roche Diagnostics, 1:2500 in the blocking buffer) at 4^{o}C for 12h. After washing and signal amplification with TSA-plus reagent (Perkin-Elmer, Waltham, MA) for 10min according to the manufacturer's instruction, the sections were incubated at 4^{o}C for 12h with anti-Dinitrophenyl-KLH antibody conjugated with Alexa488 (Molecular probes, Carlsbad, CA, 1:2000) and anti-DIG antibody conjugated with alkaline phosphatase (Roche Diagnostics, 1:2000) in the blocking buffer. After development of fluorescence, sections were cover-slipped with CC/Mount (Diagnostic Biosystems, Pleasanton, CA), and were examined under a BX51 microscope (Olympus, Tokyo, Japan) equipped with appropriate filter sets and coupled DP30BW camera (Olympus).

Microphotographed digital images were analyzed with the Metamorph software (Molecular Devices, Downingtown, PA) and the Photoshop CS4 software (Adobe Systems, Tokyo, Japan). Aside from the double color ISH, frozen sections of amphioxus heads were stained with toluidine blue to confirm exact loci of hybridization positive cells, in particular, in the anterior nerve cord. The PFA-fixed heads were embedded in OCT compound and frozen at -80^{o}C, cut at 25 μm, and stained with 0.1% toluidine blue-O (Sigma, Tokyo, Japan).

In addition, the distribution of neurons in the nerve cord was confirmed by use of a BrainStain Imaging Kit (Molecular Probes).

5-3 Results

5-3-1 Gene expression profiles of amphioxus GPH subunits

Gene expression profiles of *GPA2, GPA2LP* and *GPB5* in adult amphioxus tissues were examined with RT-PCR of first strand cDNAs from the head region, skin, gills including endostyle, muscle, testes, and ovaries (Fig. 5-1). Gene expression of *GPA2* was confirmed in the head region, skin, gills and endostyle, muscle, testes, and ovaries, while that of *GPB5* was detected in the head region, gills and endostyle, muscle, testes, and ovaries. Transcripts of *GPA2LP* gene was detected only in the testes.

5-3-2 Expression of amphioxus GPH subunit genes in the anterior nerve cord

Since considerable expression of *GPA2* and *GPB5* genes was detected in the head region by RT-PCR, this region was examined by WISH. Positive ISH signals were found in the nerve cord, which evidently showed gene expression of *GPA2* and *GPB5*, but not that of *GPA2LP* (Fig. 5-2B, C). To clarify the localization of ISH positive cells, the organization of the anterior part of nerve cord and segment numbers of myomeres are shown in a toluidine blue-stained horizontal section of the head (Fig. 5-2A). ISH signals for *GPA2* and *GPB5* mRNAs are localized in several cells in the anterior region of nerve cord that is correspondent to the segments of myomeres from numbers 3 to 8. ISH signals were not observed in other parts of the nerve cord. The sense probe did not yield any ISH signals.

ISH positive cells are longitudinally distributed at six locations in the anterior nerve cord. When observed transversally, all *GPA2* and *GPB5* positive cells are medially scattered

along the dorsal part of central canal in the nerve cord (Fig. 5-3). A transversal view shows that several *GPA2* or *GPB5* expressing cells in discrete loci form small clusters, although the numbers of cells are different among loci. As is shown in Fig. 5-4, cell bodies of most neurons were located close to the central canal, so that *GPA2* and *GPB5* expressing cells are considered to be neurons.

5-3-3 Distribution of amphioxus GPH subunits in the body

ISH signals for expression of *GPB5* gene were observed in the gills and the ovary. In the gills, the signals were located in the glandular atrial cells (Fig. 5-5A, B), while, in the ovary, signals for GPB5 mRNA were detected in small oocytes. ISH signals were not detected in larger matured oocytes (Fig. 5-5C, D). It is noteworthy that expression of *GPB5* gene was not observed in Hatschek's pit. GPA2 mRNA was not detected in the gills and the ovary likewise Hatschek's pit.

5-3-4 Coexpression of amphioxus GPH subunit genes in the anterior nerve cord

As abovementioned, cells in the dorsal part of the anterior nerve cord showed ISH positive signals for expression of *ampGPA2* and *ampGPB5* genes, so that double color ISH was carried out to confirm whether the same cells express two genes. As is clearly shown in Fig. 5-6, GPA2 and GPB5 mRNAs colocalize in the same cells in the nerve cord. In this microphotograph, GPA2 mRNA is shown in green (Fig. 5-6A), and *GPB5* in red (Fig. 5-6B). When two photographs were merged, they yielded yellowish color, indicating that exactly the same cells coexpress *GPA2* and *GPB5* genes. Such expression patterns were also seen in other tissue sections of the nerve cord.

5-4 Discussion

Most important findings in this study was the localization of *GPA2* and *GPB5* gene-expressing cells in the anterior nerve cord of amphioxus but not in Hatschek's pit, and the evidence for coexpression of these two genes in the same cells.

5-4-1 Colocalization of GPA2 and GPB5 mRNAs in the same neurons

The primary questions in my thesis are the origin of pituitary GTHs and the evolution of the HPA. The results in this chapter yielded valuable information to answer the former question, that is, the same cells in the anterior nerve cord coexpress *GPA2* and *GPB5* genes.

An approach by the RT-PCR and then WISH provided important information that *GPA2* and/or *GPB5* expressing cells are localized in the distinctive region of the anterior nerve cord, that is, the anterior region of nerve cord corresponding to the segments of myomeres from numbers 3 to 8. There were an abundance of papers which reported the presence of hypothalamic neurons that synthesize pituitary hormones including gonadotropins (Civelli et al., 1982; Freeman et al., 2000; Parhar et al., 2003). Hence, the present finding in amphioxus is not surprising, but rather provides evidence for evolutionary origin of central neurons which secrete pituitary hormones.

In combination with the findings in Chapter 3 that showed ability of dimerization between ampGPA2 and ampGPB5, coexpression of *GPA2* and *GPB5* in the same cells supports the idea that amphioxus neurons in the dorsal part of anterior nerve cord secrete thyrostimulin as heterodimeric GPH, although, to further establish this idea, various lines of evidence is required, such as assembly of ampGPA2 and ampGPB5 to form heterodimer in the endoplasmic reticulum like that of pituitary GTHs (Xing et al., 2004).

5-4-2 Gene expression profile of amphioxus GPH subunits

Analyses by RT-PCR showed the presence of GPA2 and GPB5 mRNAs in various

tissues, such as the head, gills including endostyles, muscles, and gonads in amphioxus. Similarly, in vertebrates, *GPA2* and *GPB5* genes are expressed in various tissues (Nakabayashi et al., 2002; Hsu et al., 2002; Okada et al., 2006). Although ISH with *GPB5* specific probes showed positive signals in the gills and gonads, *GPA2* specific probe did not provide any positive signals in these tissues. This discrepancy may be explained by differences in the levels of gene expression between two genes, or among tissues.

5-5 Conclusion

Transcripts of genes for three GPH subunits were localized in adult amphioxus. GPA2 and GPB5 mRNAs were detected in various tissues including the anterior nerve cord. Double color ISH showed colocalization of two subunit mRNAs in the same neurons in the dorsal part of anterior nerve cord. In combination with findings in the previous chapter that recombinant GPA2 and GPB5 probably form a heterodimer, the results in this chapter suggest that amphioxus GPA2 and GPB5 form thyrostimulin, which is dimeric, and secreted from neurons in the nerve cord.

Table 5-1. Primer sequences used for RT-PCR of *ampGPA2*, *ampGPA2LP*, and *ampGPB5*.

Primer name	Sequence (5' to 3')	Experiment
GPA2-F	ATCATAGAACATTCAGAAGCCAGTC	RT-PCR
GPA2-R	TCTTCTTCAAGGTTCTTCTTGCCAA	RT-PCR
GPA2LP-F	TGTTTATGGCGTTTTGTAGCG	RT-PCR
GPA2LP-R	TACATCACGGGTTGTAGTCAC	RT-PCR
GPB5-F	GGTGTCTTGTGACCATATGCATCTG	RT-PCR
GPB5-R	TACCATCCACAATCGTTTTCCAATG	RT-PCR

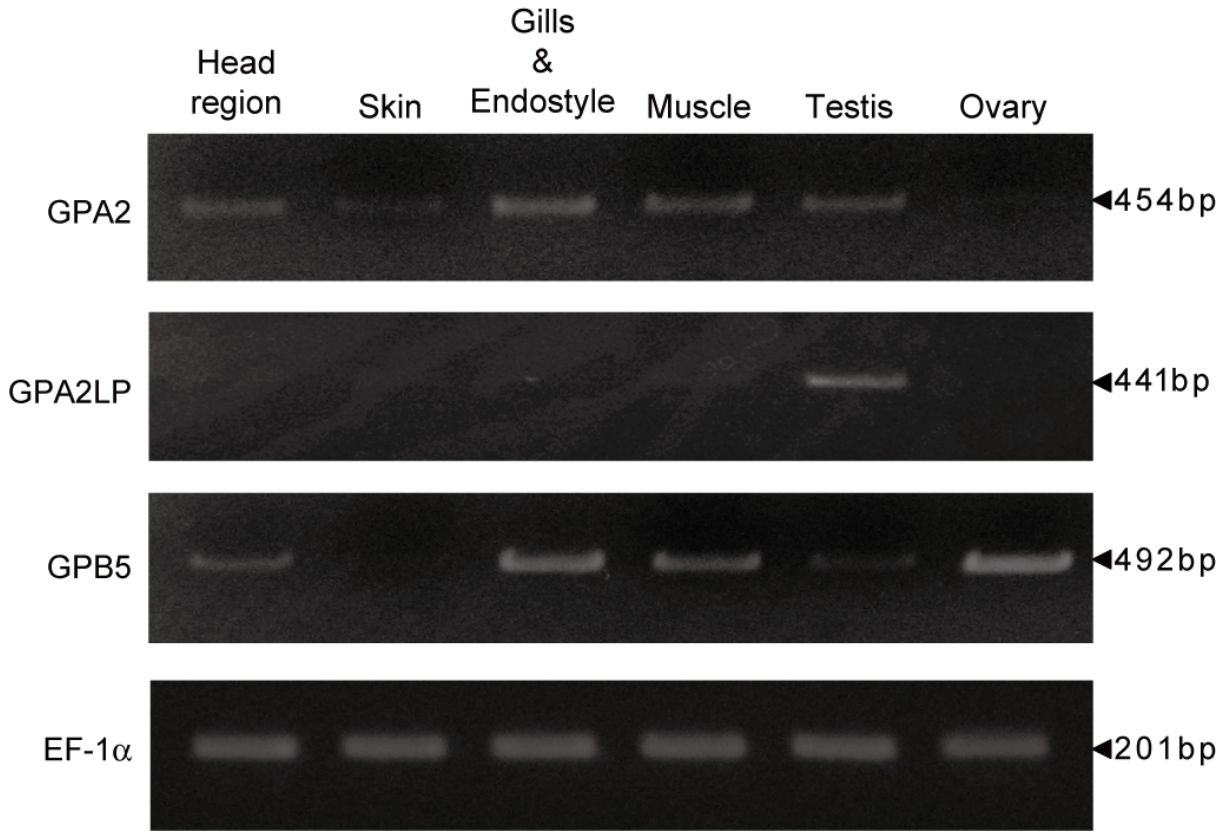

Figure 5-1. Gene expression profiles of *ampGPA2*, *ampGPA2LP*, and *ampGPB5* in various tissues of amphioxus. Total RNAs were extracted from the head region, skin, gills including endostyle, muscle, testis and ovary of animals collected in breeding season. First strand cDNAs were amplified using the specific primers for each gene. EF-1α mRNA was used as the internal control.

Fig. 5-2.

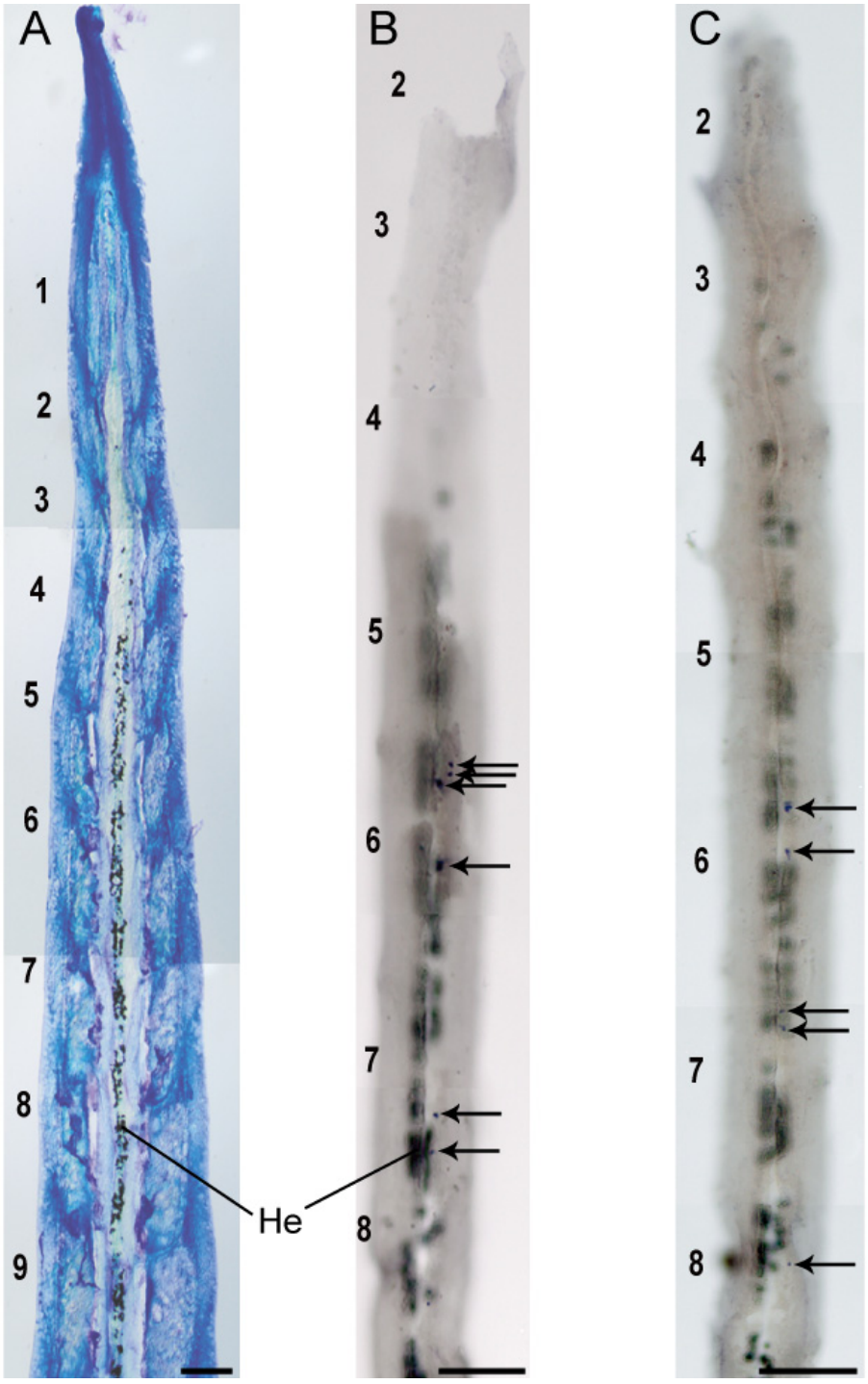

Figure 5-2. ISH analyses showing localization of gene transcripts for *ampGPA2* and *ampGPB5* in the nerve cord of amphioxus. (A) Horizontal section of the anterior region of nerve cord stained with toluidine blue. Note that the Hesse organs containing pigment cells are shown as dark spots behind myomere 3 on the midline of the nerve cord. Signals for ampGPA2 mRNA (B) and ampGPB5 mRNA (C) are localized in the dorsal area of anterior nerve cord. Arrows indicate ISH positive, and numbers in the panels indicate the myomere numbers. He, Hesse organs. Scale bars, 100 μm.

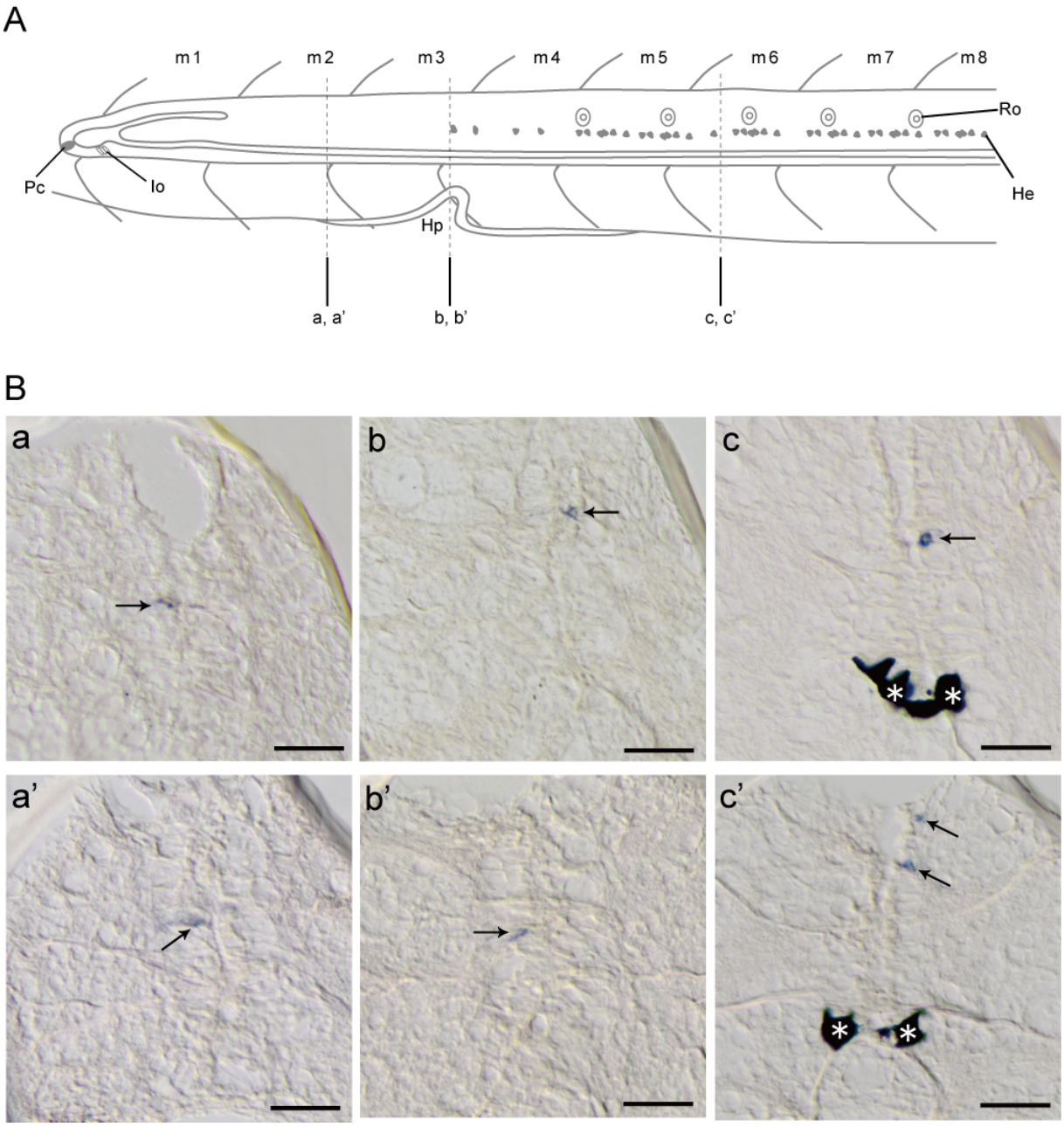

Figure 5-3. Expression of *ampGPA2* and *ampGPB5* in the anterior part of the nerve cord of amphioxus. (A) Schematic lateral view of the anterior nerve cord. Transverse frozen sections shown in the lower panel were cut at the position indicated by the vertical bars. (B) Expression of *ampGPA2* in neurons which locate in the middle (a) and dorsal (b, c) areas along the central canal. Expression of *ampGPB5* was also localized in similar areas (a'-c'). Arrows indicate ISH positive nerve cells. White asterisks indicate Hesse organs (c, c'). He, Hesse organ; Hp, Hatschek's pit; Io, infundibular organ; m, myomere; Pc, pigment cells of frontal eye; Ro, cell of Rohde. Scale bars, 20 μm.

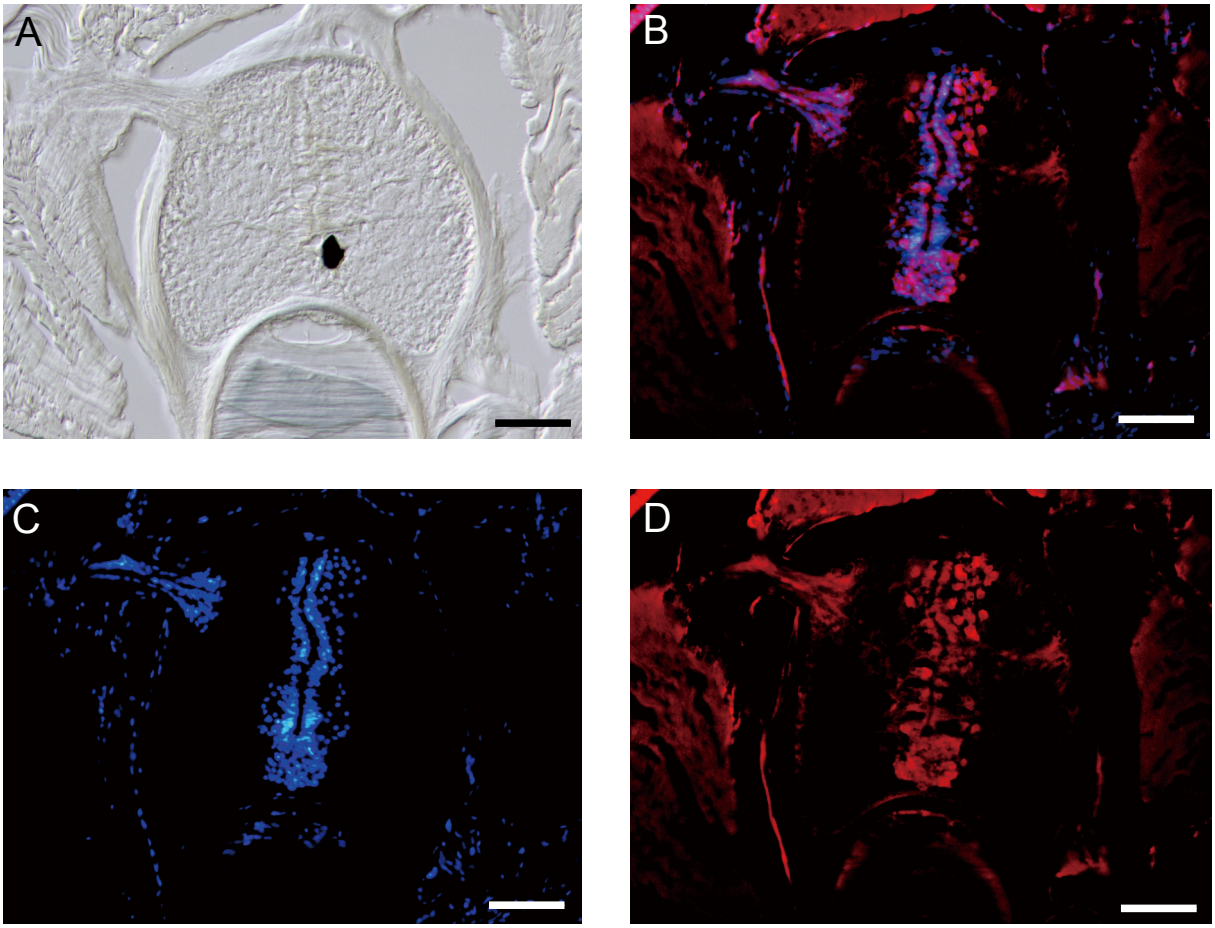

Figure 5-4. Architecture of the nerve cord shown by DAPI and Neuro Trace 530/612 fluorescent staining. (A) Unstained transverse section of the nerve cord at the level of myomere 5. Note the arrangement of the central canal, Hesse organ with pigment cells, and the projection of the nerve root from the left side of nerve cord to muscle. (B) Merged image of DAPI and Neuro Trace 530/612 stained cells. Nuclei are labeled blue as in C, while Nissl bodies are stained red as in D. Images of C and D are combined. Note that central neurons are distributed around the central canal. (C) Staining of nuclei with blue DAPI. (D) Staining of Nissl bodies with red Neuro Trace 530/615. Scale bars, 25 μm.

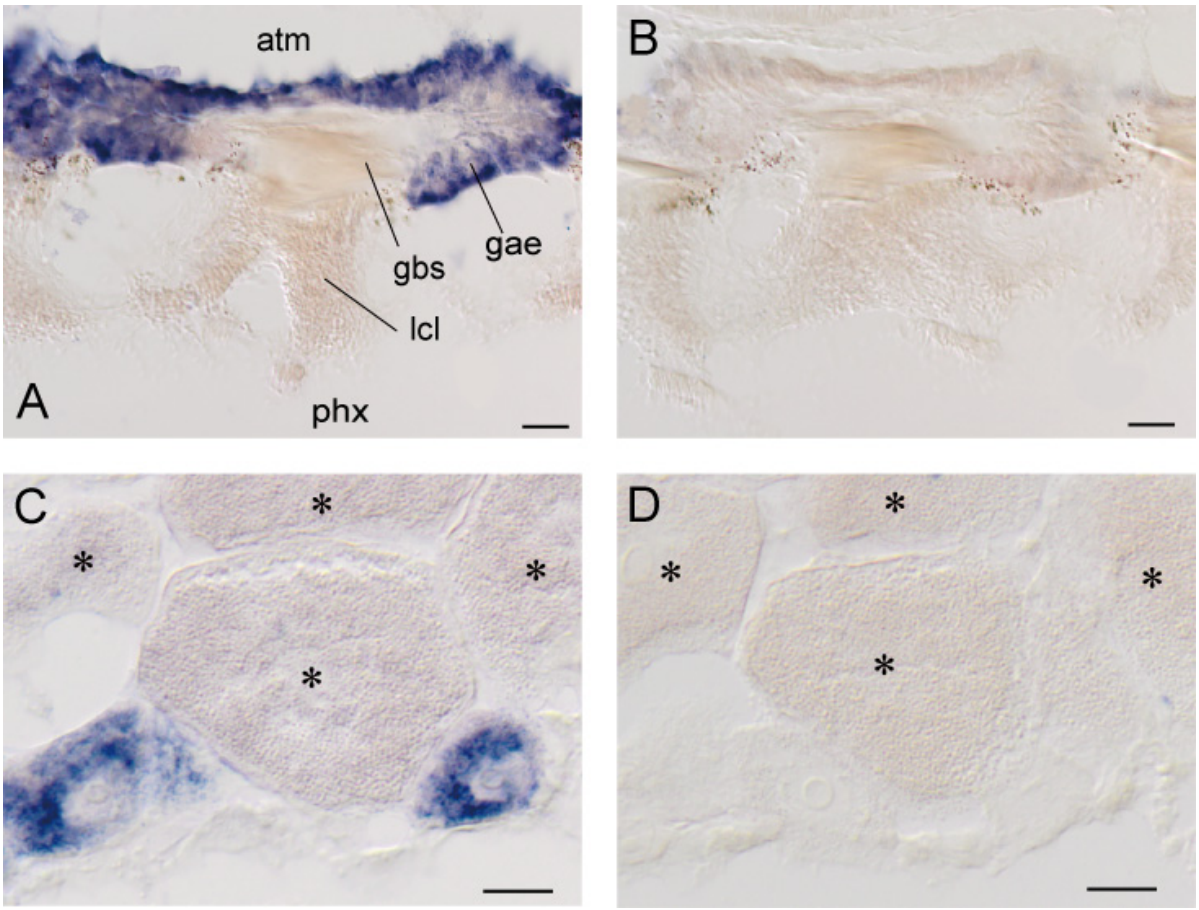

Figure 5-5. Expressions of *ampGPB5* in the gill (A, B) and the ovary (C, D) of amphioxus. Note ISH positive cells for ampGPB5 mRNA in glandular cells of gills (A), and small developing oocytes (C). No signals exist in the gills and the ovary when a sense probe was used in adjacent sections (B, D). Asterisks in C and D indicate vitellogenic oocytes. atm, atrium; gae, glandular cells; gbs, gill bar skeleton; lcl, lateral cells; phx, pharynx. Scale bars, 20 μm.

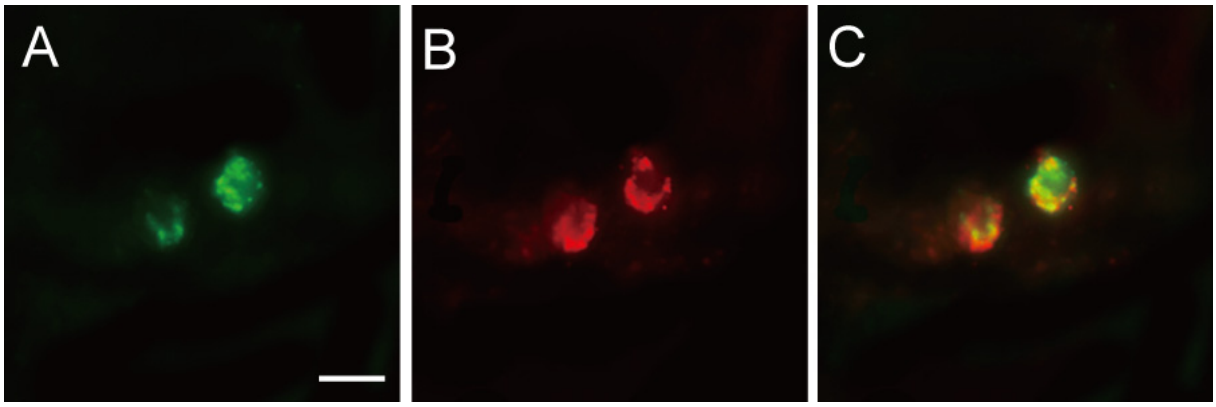

Figure 5-6. Double color immunofluorescent ISH for expression of *ampGPA2* and *ampGPB5* in the anterior nerve cord. Positive cells for expression of *ampGPA2* are labeled green (A), and those of *ampGPB5* are stained red (B). Mergence of these images (C) describes co-expression of *ampGPA2* and *ampGPB5* in the same cells. Scale bar, 10 μm.

Chapter 6

Expression of Genes Encoding Neurohypophysial Hormone and Receptors for Hypothalamic Hormone and Sex Steroid in the Head of Amphioxus

6-1 Introduction

In vertebrates, the hypothalamus-pituitary-target organ axes are important endocrine systems that regulate various physiological functions. Hypothalamic neurosecretory neurons that control pituitary functions release neurohormones from axon terminals in the median eminence into blood capillaries of the portal vessels. These neurohormones, mostly peptidergic, are carried to the adenohypophysis via the portal vessels, and regulate syntheses and releases of proteinaceous pituitary hormones like GH, PRL, GTHs, TSH and ACTH. Neurohormones either stimulate or inhibit particular pituitary cells to release their own hormone into general circulation, so that functions of target organs, such as adrenal grand, thyroid gland, and gonads, are regulated adequately. On the other hand, classical hypothalamic neurosecretory neurons that secrete neurohypophysial hormones send their axons mainly to the pars nervosa, and release their hormones into general circulation. These hormones are also involved in controls of physiological and reproductive function.

In my thesis, I tried to find homologues of pituitary hormones in Hatschek's pit, which has been considered as a homologue of the pituitary gland, however, analysis of gene transcripts and proteins of Hatschek's pit did not detect expression of pituitary hormone-related molecules. Instead, I obtained thyrostimulin subunit genes and found that *GPA2* and *GPB5* mRNAs were detected in the anterior nerve cord. This finding raised a question whether nerve cells which secrete thyrostimulin have any neuroendocrine functions in connection with other neurons.

Here, I focused on reproductive function, because the recent survey of entire genome sequences of *B. floridae* showed the presence of homologous genes encoding vasotocin-like peptide (ampVT) and gonadotropin-releasing hormone receptor (GnRHR) (Fig.6-1) (Holland et al., 2008). These two substances are involved in the control of reproduction in vertebrates, as well as pituitary GTHs homologous to thyrostimulin. In addition, Gwee et al. (2009)

reported the analysis of a locus of vasotocin gene in *B. floridae*, and demonstrated that amphioxus has one homologous gene of a neuroypophysial hormone gene; and Tello and Shewood (2009) cloned four cDNAs encoding amphioxus GnRHR(ampGnRHR) that are referred to as ampGnRHR 1 to 4 from *B. floridae* and compared their binding activities to several types of GnRHs.

To approach the above question, I adopted an ISH technique to clarify localization of neurons which express genes encoding ampVT and ampGnRHR. Nucleotide sequence of cDNA for ampVT was obtained from *B. belcheri*. Since sex steroid hormones have crucial roles in reproduction in vertebrates, distribution of mRNAs for estrogen receptor (ampER) and steroid receptor (ampSR) were also examined in the head region of amphioxus. Then, distribution of hybridization signals for these signal substances were compared with those for ampGPA2 and ampGPB5.

6-2 Materials & Methods

6-2-1 Cloing of ampVT *cDNA*

The protocol of ampVT cDNA cloning followed that described in Chapter 3. The primers used for cloning of *ampVT* for partial fragment are:

5’-CATCATCAACTGTCCCCG-3’, and 5’-CAGGTTCTCCATGGAGCA-3’,

and for full length cloning,

VtGSP-R 5’-TGCAGCACGTGGTGGGCCCGATA-3’ and

VtGSP-R_Nest 5’-GTCCGCAACGAGGGCACTCTCT-3’ for 5’-RACE

VtGSP-F 5’-GCGGGAAGAGAGCCCTGGAAAC-3’ and

VtGSP-F_Nest 5’-GTCGCTGTCGGGCCAGTGTATC-3’ for 3’ RACE

6-2-2 ISH for AmpVT, AmpGnRHRs, AmpER, *and* AmpSR

To construct a molecular probe for ISH of ampVT mRNA, *ampVT* cDNA was used as the template, and amplified by PCR with a set of two primers: vasotocin-F 5'-CAGCTGCGTAGGTTTTTGG-3' and vasotocin-R 5'- CGACGCATACAATATTCATCT-3' (nt 13-959; 947 bp). The PCR product was purified with a QIAX II gel extraction kit (QIAGEN) and subcloned into a pCR4 TOPO plasmid vector (Invitrogen). cDNA clones of *ampGnRHR1, 2, 3, and 4* were kindly provided by Dr. Nancy M. Sherwood (University of Victoria). These cDNAs were cloned from *B. floridae* (Genbank accession numbers: GnRHR1, EU433377; GnRHR2, EU433378; GnRHR3, EU433380; and GnRHR4, FJ426561) and ligated in pcDNA 3.1 plasmid vector (Invitrogen). Insert sequences were excised and subcloned into pCR4 TOPO plasmid vector by PCR amplification using gene specific primers for each gene, and used to construct molecular probes. Full length cDNAs for *ampER* and a*mpSR* ligated in pcCR Blunt II from *B. belcheri* (Genbank accession numbers are ampER, AB10027; ampSR, AB510028) were kindly provided by Dr. Y. Katsu (Hokkaido University). Construction of molecular probes was conducted as described in Chapter 3. The head regions of mature animals were fixed in 4% PFA and embedded in TissueTek OCT compound (Ted Pella), and cross-sectioned by a cryostat at 20 μm. ISH was performed as described in Chapter 5.

6-3 Results

6-3-1 Cloning of ampVT *cDNA*

Full length cDNA encoding ampVT precursor was cloned from *B. belcheri* using a combination of RT-PCR and RACE techniques (Fig. 6-2A). The cDNA is 989 bp in length and contains an open reading frame of 501 bp (167 amino acids). AmpVT precursor

consists of signal peptide, nonapeptide hormone, neurophysin and copeptin, although the border of neurophysin and copeptin is obscure. A similarity of vasotocin precursors between *B. belcheri* and *B. floridae* (protein ID in the genome database is 84802, Gwee et al., 2009) is 90% in amino acids and 91% in nucleotides. Ser^4 in *B. floridae* ampVT is replaced with Ile in *B. belcheri* ampVT. Prohormone of this nonapeptide hormone is connected to neurophysin with a typical tripeptide sequences (G-K-R) that is known as a signal for proteolytic processing and C-terminal amidation of peptide hormones. Copeptin of amphioxus lacks a leucin-rich core (Fig. 6-2B).

6-3-2 Distribution of ampVT mRNA in the nerve cord

ISH signals for ampVT mRNA were found in the anterior part of nerve cord, indicating evident expression of *ampVT* in this region. Sense probes did not yield any ISH signals. Exact localization of ISH positive cells in the anterior nerve cord were identified in comparison with segment numbers of myomeres (Fig. 6-3A).

According to their sizes and location, ISH positive cells are divided into two populations as the anterior parvocellular group and the posterior magnocelluar group (Fig. 6-3B). The anterior group is composed of small cells with diameters of ca. 10 μm. These cells, which form a small cell mass ranging over 50 μm, locate bilaterally along the ventral half of central canal in the rostral part of nerve cord (Fig. 6-3B-a, B-c). This locus corresponds to the infundibular organ.

The posterior group was localized in the medial region of anterior nerve cord, rostro-caudal location of which corresponds to the segments of myomeres 3 to 6. The heavily stained bipolar cells with longitudinal lengths over 20 μm lie across the rather dorsal part of central canal (Fig. 6-3B-b, d, e). Neuronal processes arising from somata of these cells run laterad and bifurcate in neuropiles (Fig. 6-3B-d).

6-3-3 Distribution of mRNA encoding ampGnRHRs in theanterior nerve cord

Among mRNAs encoding four types of GnRHR, ISH signals showing expression of *ampGnRHR1* and *ampGnRHR2* were observed in the anterior nerve cord. Any signals were not obtained for *ampGnRHR3* and *ampGnRHR4*. The organization of the anterior nerve cord and segment numbers of myomeres are shown in the diagram (Fig. 6-4A, 6-5A). ISH signals for ampGnRHR1 and ampGnRHR2 mRNAs were found in several cells in the anterior region of nerve cord that is correspondent to the segment of myomeres from number 1 to number 8 (Fig. 6-4B-b, d, e). The sense probe did not yield any ISH signals.

When observed transversally, *ampGnRHR1* and *ampGnRHR2* positive cells are medially scattered along the dorsal and middle parts of the central canal in the nerve cord (Fig. 6-4B, 6-5B). A mid-sagittal view shows that several *ampGnRHR1* and *ampGnRHR2* expressing cells in discrete loci form small clusters, although the numbers of cells are different among loci.

6-3-4 Distribution of steroid receptor genes in the anterior nerve cord

ISH signals showing expression of *ampER* were also found in the anterior nerve cord. The organization of the anterior nerve cord and segment numbers of myomeres are shown in the diagram (Fig. 6-6A). ISH signals for ampER mRNA were localized in several cells in the cerebral portion of the apical part of the nerve cord and the anterior region of nerve cord that is correspondent to the segment of myomeres from number 1 to number 8 (Fig. 6-6B). The sense probe did not yield any ISH signals.

Positive signals are shown in the cerebral portion in transverse sections (Fig. 6-6Ba), and signal intensity is stronger at its dorsal area than the ventral area. The labeling of *ampER* transcripts gradually diminished toward the posterior region and disappeared behind

the starting point of myomere 1 (Fig. 6-6B-d).

Another population of *ampER*-positive cells was found in more posterior region of the nerve cord, around myomeres 3 and 4 (Fig. 6-6B-b). The region between myomeres 3 and 4 is close to Hatschek's pit. In this region, ISH signals were seen in the dorso-ventrally middle region of the nerve cord along the central canal. Transversely, faint *ER*-positive cells were located in the middle region face to the ependymal layer of the central canal. More intense signals were detected in the the floor of the central canal (Fig. 6-6B-d). Diffused weak signals for *ampER* were shown in the Hatschek's pit (Fig. 6-6B-e). No positive signals are obtained for *ampSR*.

6-4 Discussion

The present ISH analysis of the nerve cord of amphioxus evinced expression of genes encoding ampVT, ampGnRHR 1 and 2, and ampER. Cells that express these genes are distributed in the discrete regions in the anterior part of nerve cord, where *ampGPA2* and *ampGPB5* are coexpressed in the same cells (Chapter 5).

6-4-1 Reliability of present ISH

In the present study, most hybridization probes except those for GnRHR were prepared by use of cDNAs obtained from *B. belcheri*. Further, RNA probes are well known to be highly sensitive to mismatching. Thus, my present protocol ensured that positive ISH signals were derived from highly specific matching between the probe and particular mRNA. An arising question here is the lack of ISH signals for mRNAs encoding GnRHR3 and 4, although positive signals could be obtained for GnRHR1 and 2 mRNAs. One of possible reasons for this difference is the use of cDNAs from *B. floridae* in the different species, if the

sequences of *ampGnRH3* and *4* are not highly identical. Otherwise genes for GnRH3 and 4 are actually not expressed in the anterior nerve cord of *B. belcheri*.

6-4-2 Distribution of ampVT transcripts in the anterior nerve cord

My ISH study showed the presence of two populations of *ampVT*-positive cells, one in the cerebral part and the other in the anterior nerve cord. In the cerebral part, perikarya of large neurons which localize in the caudal part of infundibular organs contain numerous electron-dense granules (Meves, 1973). The localization and the shape of these neurosecretory cells and the *ampVT*-positive cells are similar in the cerebral part. It is suggested that ampVT is synthesized in neurosecretory cells in the cerebral part which corresponds to the brain of vertebrates.

In the second population, these cells that locate between myomere 3 and 6 are neurons with perikarya crossing the central canal and projecting processes to both sides of the nerve cord. Cells showing this structure were reported by microscopic investigation and called commissural cells (Franz, 1924; Bone, 1960; Meves, 1973). According to the report by Bone (1960), commissural cells at the middle of the central canal project axons into the bundle compartments. Further, commissural cells similar to *ampVT* positive cells were reported to be immunoreactive to anti-FMRFamides, gamma-aminobutyric acid (GABA), and NPY (Pestarino and Lucaroni, 1996; Anadon et al., 1998; Castro et al., 2003). These reports suggest that ISH positive *ampVT* cells correspond to commissural cells.

The previous immunohistochemical study in which anti-vasotocin and anti-vasopressin were used showed that immunoreactive cells were distributed in the lateral central canal in the spinal cord but not in the cerebral part of the nerve cord (Uemura et al., 1994). In my study, ISH positive ampVT cells are distributed in the cerebral part and also in the spinal cord. It is therefore possibile that ampVT is synthesized in neurosecretory cells in the cerebral part and

then transported in projecting fibers to the caudal part of the nerve cord.

The presence of homologous peptides to neurohypophysial hormones in invertebrates suggests that the vasopressin-oxytocin family originated before the proto- and deuterostomia (Cruz et al., 1987; Proux et al., 1987; Reich et al., 1992; Salzet et al., 1993; Oumi et al., 1994; Takuwa-Kuroda et al., 2003; Li et al., 2008; Stafflinger et al., 2008). Most of them were purified from central nervous tissues, and the transcripts were distributed in neurons, e.g., annetocin in the subesophageal ganglia in earthworms (Satake et al., 1999) and octopressin in the brain of octopus (Takuwa-Kuroda et al., 2003).

When these findings are compared with information on vertebrates, distributional patterns of vasopressin-oxytocin family hormones differ between invertebrates and vertebrates. Hormones are widely scattered in various neurons in invertebrates, whereas integrated in neuroendocrine cells in the brains of vertebrates. Distribution of ampVT in the rostral area of the nerve cord may indicate a primordial form of vasopressin-oxytocin family hormones before emergence of the brain and the pituitary.

6-4-3 Distribution of ampGnRHRs *in the anterior nerve cord*

My present ISH showed that, among four *ampGnRHRs* genes, *ampGnRHR1* and *ampGnRHR2* were transcribed in the anterior nerve cord, in particular in its caudal region at the level of myomere 1. According to Ekhart et al (2003), there is a cluster of medium-sized neurons in the lateral periventricular grey that flanks on the ventricle. The position of *ampGnRHR1* and *ampGnRHR2* positive cells correspond to these cells in this cluster.

Discontinuous dot-like ISH signals for expression of *ampGnRHR1* and *ampGnRHR2* in small cells caudal to myomere 2 partially overlap with the region where expressions of *ampGPA2* and *ampGPB5* genes were investigated in Chapter 3. Interesting question remained to be clarified is a relation between ampGnRHRs and ampGPA2/ampGPB5,

particularly their colocalization.

GnRH and GnRHR are widely present from invertebrates through vertebrates (Kah et al., 2007). Tissue distributions of GnRHRs, as well as GnRHs, in tunicate and octopus demonstrated that their genes were expressed in various tissues (Kusakabe et al., 2003; Kanda et al., 2006). These results suggest multifunctional roles of GnRH as a neurotransmitter, neuromodulator, and hormone-like factor (Kah et al., 2007). Identification of GnRH and confirmation of its distribution in amphioxus, as well as understanding of interaction with GnRHRs, will necessary to reveal the function of amphioxus GnRH.

6-4-4 Distribution of ampER *in the anterior part of amphioxus*

Transcripts of *ampER* were detected bilaterally in the anterolateral periventricular cell groups. This distribution is almost the same with the expression pattern of amphioxus gene homologous to *Period (Per)*, which is related to the circadian system in vertebrates (Schomerus et al., 2008). They suggested homology between these cell groups and the vertebrate suprachiasmatic nuclei (SCN) because of the similarity of shape between *Per*-positive cell groups of amphioxus and SCN of vertebrates. This idea is consistent with another report that GABA, a typical neurotransmitter of the SCN in mammals (Moore and Speh, 1993), is present in these cell groups of the amphioxus nerve cord (Anadon et al., 1998). If this idea is true, it would be worthwhile to demonstrate ERs in the SCN of vertebrates, such as mammalian (Shughrue et al., 1997) and teleost species (Hawkins et al., 2000).

ISH positive *ampER* cells also appeared in more caudal region of the amphioxus nerve cord at the level of myomere number 3 and 4. Most of positive signals were diffusely distributed along the central canal, and rather intense signals were detected at the ventral region of the central canal. The most conspicuous feature of this area is the nucleus of Rhode (Ekhart et al., 2003). This is an agglomeration of relatively large cells that surround

the ventral expansion of the central canal. Interestingly, rostrocaudal extension of the nucleus of Rhode coincides with that of the columnar epithelium of the wheel organ and Hatschek's pit (Ekhart et al., 2003). The *ampER*-positive cells are located in this cell group. However, the number of positive cells is quite few, when compared with the nucleus of Rhode. This suggests that, not all, but several nuclei of Rhode express *ampER*.

Transcripts of *ampER* were also detected in Hatschek's pit. This is the first finding of the presence of endocrine-related genes in Hatschek's pit by ISH. Recently, the presence of NPY immunoreactive cells was reported in Hatschek's pit (Castro et al., 2003). The slender shaped NPY immunoreactive cells were more like type 1 cells than type 2 cells, and it is rare in the epithelium covering the bottom of the pit. This pattern is similar to the distribution of *ampER*, although NPY immunoreactive cells appear intensely to delineate the cells and observed in their apical regions.

Katsu et al. (2009) demonstrated that ampER does not bind estrogen and represses activation of ampSR which has a potential to bind estrogen, so that estrogen mediated transcriptional regulation by direct binding with estrogen is not conceivable as a function of ampER. However, my present study did not detect expression of ampSR gene in any regions, despite constitutive expression of ampER. Further investigation are required to clarify ampSR gene expression profiles associated with seasonal, physiological, and environmental changes.

6-5 Conclusion

In this chapter, ISH technique was applied to clarify localization of neurons which express genes encoding ampVT and ampGnRHR, as well as steroid receptors, in relation to the distribution of ampGPA2 and ampGPB5 expressing cells. The findings in this chapter are summarized in Fig. 6-7. *AmpVt* transcripts are demonstrated in two obvious populations

of positive cells, presumable neurosecretory cells located in the caudal part of the cerebral vesicle, and commissural cells located in the nerve cord lying in between myomeres 3 and 6. Among four identified *ampGnRHRs*, two a*mpGnRHRs* showed their gene expression in a pair of cells at both sides of the central canal located at the level of myomere 1 and more rostral part of the anterior nerve cord. *AmpER* transcripts were distributed in the rostral region of the nerve cord and more caudal part, which correspond to the region near the Hatschek's pit. Furthermore, its gene expression was found in the cells of Hatschek's pit.

In conclusion, the presence of amphioxus homologous genes of neurohypophysial hormone, hypothalamic hormone receptors, and steroid hormone receptor are distributed in the anterior part of the nerve cord, especially in its rostral part.

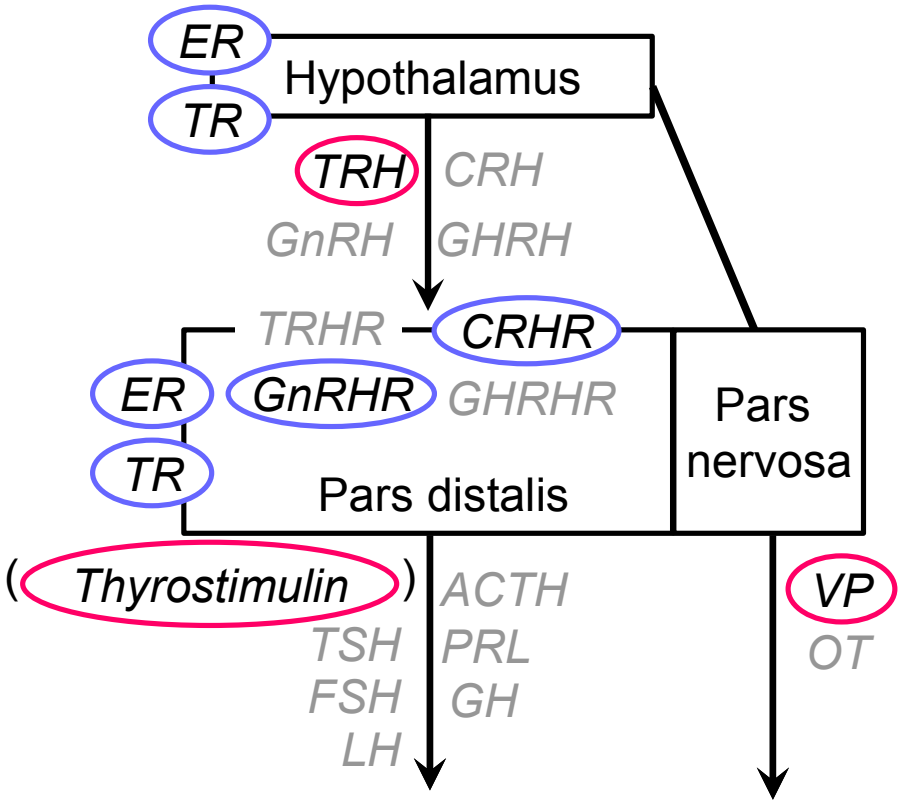

Figure 6-1. Hormones and their receptors in the hypothalamus and the pituitary in vertebrates. Analyses of the genome database of *Branchiostoma floridae* suggest that amphioxus have thyrostimulin subunits, thyroid hormone releasing hormone (TRH) and vasopressin (VP) shown by dark letters circled with pink line. Receptors to be identified in amphioxus are corticotropin releasing hormone receptor (CRHR), estrogen receptor (ER), gonadotropin-releasing hormone receptor (GnRHR), thyroid hormone receptor (TR) indicated by dark letters circled with blue line. Pale letters indicate vertebrate-specific hormones and receptors probably not present in amphioxus. Thick lines indicate the cascade of hormonal action.

Fig. 6-2.

A

```
  1 GACTCCACGTGCGCAGCTGCGTAGGTTTTTGGATATGAGGTTGTAGTGCTCGACCTCGACTCATCCTGTTCCGTCCGTCATTGTCCATTT 90
 91 CTTCTGACGTCTTCGCTGAGACAACCTCTACATAAGGCGCGCCGTCTCTGCTCAGTTCTCACCTACCAGAGCGGAGAGACATCAGCGGAT 180
181 AGAATGTGGCGGTTTGGGATGGTCCTGTCCGCGGTGCTGATACTGGTGGTGGTGGCGTCCACCATAGGACGGTCTGCCGGCTGCTACATC 270
  1    M  W  R  F  G  M  V  L  S  A  V  L  I  L  V  V  V  A  S  T  I  G  R  S  A  G  C  Y  I    29

271 AGCAACTGTCCCCGGGGCGGGAAGAGAGCCCTGGAAACCAGGAGCGGGCCTGCTGGGAGAGAGTGCCCTCGTTGCGGACCGTCGCTGTCG 360
 30 S  N  C  P  R  G  G  K  R  A  L  E  T  R  S  G  P  A  G  R  E  C  P  R  C  G  P  S  L  S    59

361 GGTCAGTGTATCGGGCCCACCACGTGCTGCAGCCCGCTGGCCGGCTGTACCCGCAGCCTGAGCGTGGCCCTGGAGTGCTCCATGGAGAAC 450
 60 G  Q  C  I  G  P  T  T  C  C  S  P  L  A  G  C  T  R  S  L  S  V  A  L  E  C  S  M  E  N    89

451 CTGGTGCCCGTGCCGTGCAGACTGGGCGGCCCCTCCTGTACCCTGCCCGGCCAACAGACGGGTACCTGCGTTGGGGAAGGGATGTGCTGC 540
 90 L  V  P  V  P  C  R  L  G  G  P  S  C  T  L  P  G  Q  Q  T  G  T  C  V  G  E  G  M  C  C   119

541 ATGGACGGAGAAAAATGCAGTCTGTCTTCAGAATGCTCCGCGAGGAACGCTGAAGACGAAGAAGCAAGAAGGCAAGACCGGCGGTCGGTG 630
120 M  D  G  E  K  C  S  L  S  S  E  C  S  A  R  N  A  E  D  E  E  A  R  R  Q  D  R  R  S  V   149

631 CTGCCCCACTGGGTTCTGTCCGGGAACACCCCTCCGGACGTGCAGAAATGGTGGTGACGGATCGGCCGCCTTTCGTGCGTCCATATATTC 720
150 L  P  H  W  V  L  S  G  N  T  P  P  D  V  Q  K  W  W  *                                   167

721 ACTAAGAACCGCAGGGCTATTCCATTCGGAGGGGTGTAACTGACAAAAATGTTCAATACATTCTTCTTTTTCATGAGACGTCACAATCGC 810
811 GTTCGAACCAGAAGTCGGCCATGTTGGATGATAAACTGTTCTATTTTGCCAGGTTAGCATCCAACATGGCGCCGATGAGGTGACGTAGCC 900
901 ATTTTCAAGAGATCAATTGTTATGGCAACGCTGAAATAAAGGAAAAGATGAATATTGTATGCGTCGGGATAAAAAAAAAAAAAAAAAAA  989
```

B

```
                          Signal peptide          Hormone          Neurophysin
human_VP        1:-------MPDTMLPACFLGLLAFSSACYFQNCPRGGKRAMSDLE---LRQCLPCGPGGKG 50
chick_VP        1:-------MAEPSLPLSFLCLLALSSACYIQNCPRGGKRALGDTA---LRQCLPCGPGNRG 50
xenopus_VT      1:-------MPEASVPACFLCLLALSSACYIQNCPRGGKRSYPDTE---LRQCMQCGPGNRG 50
lungfish_VT     1:-------MPGTCLPLCFLCLLAFSSACYIQNCPRGGKRSFIDTE---IRQCIPCGPQNRG 50
takifugu_VT     1:-------MPQCALLLSLLGLLALSSACYIQNCPRGGKRALPETG---IRQCMSCGPRDRG 50
lamprey_VT      1:----MARCAPLTLAVSVLSLVLISSACYIQNCPRGGKRDLTDS----VRQCLPCGPGGQG 52
amphioxus_VT    1:MWRFGMVLSAVLILVVVASTIGRSAGCYISNCPRGGKRALETRSGPAGRECPRCGPSLSG 60
                                             **  ********          * *  **    *

                                                                                
human_VP       51:RCFGPSICCADELGCFVGTAEALRCQEENYLPSPCQSGQKACGS----GGRCAAFGVCCN 106
chick_VP       51:RCFGPGICCGAELGCYLGTAETRRCAEEDYMPSPCQAGGQPCGS----DGRCAANGVCCS 106
xenopus_VT     51:NCFGPNICCGEDMGCYIGTPETLRCVEENFVPSPCEAGGRPCST----GGRCAAPGICCN 106
lungfish_VT    51:RCFGPYICCGEELGCYIGTSETLRCLEENYLSSPCKAGGKLCSTN---GGQCAAPGICCT 107
takifugu_VT    51:RCFGPNICCGEALGCLMGSPETARCAGENYLLTPCQAGGRPCGSE---GGRCAVSGLCCN 107
lamprey_VT     53:RCFGPRICCGEAMGCRLGGPDVAICRAERLMPSPCESRGEPCGH----GGKCGAPGLCCS 108
amphioxus_VT   61:QCIGPTTCCSPLAGCTRSLSVALECSMENLVPVPCRLGGPSCTLPGQQTGTCVGEGMCCM 120
                   * *   **    **         *  *     **      *         *   * **

                                                  Copeptin
human_VP      107:D-ESCVTEPECREGFHRRA-RASDRSNATQLDGPAGALLLRLVQLAGAPEPFEPAQPDAY 164
chick_VP      107:A-DTCAMDAVCLEEGSEQA-EEAAEKNLTVLDGAAGDLLLRLMHLANRQQQGKQPGL--- 161
xenopus_VT    107:D-ESCSLDSACLDDESERR-RAPLEKNTTVMDGSASDFLLRLMHMANRQQQAKHQYY--- 161
lungfish_VT   108:D-ESCAMDSSCLDGDADKR-RMFPERNLTLLDGATSDFLLKLIHLANRQQQEEKHLL--- 162
takifugu_VT   108:S-ESCAVDSDCLG-ETES-----LEPGDSSADSSPTELLLRLLHMSSRGQSEY------- 153
lamprey_VT    109:S-ESCAEDASCGWEGGDSPGERPFPHSALRLQSPAAEAMLELINSNSLRD---------- 157
amphioxus_VT  121:DGEKCSLSSECSARNAEDE-EARRQDRRSVLPHWVLSGNTPPDVQKWW------------ 167
                        *     *
```

Figure 6-2. (A) Nucleotide and amino acid sequences for precursors of ampVT, and (B) multiple alignments of precursors for ampVT, vasopressin and vasotocin in various vertebrates. The conserved amino acid residues among animals are marked with asterisk. The DDBJ/EMBL/GenBank accession numbers of sequences used for analysis are as follows. human VP (NP000481); chick VP (NP990516); xenopus VT (CT025342); takifugu VT (AB297919); lamprey VT (BAA06669). VP, vasopressin; VT, vasotocin.

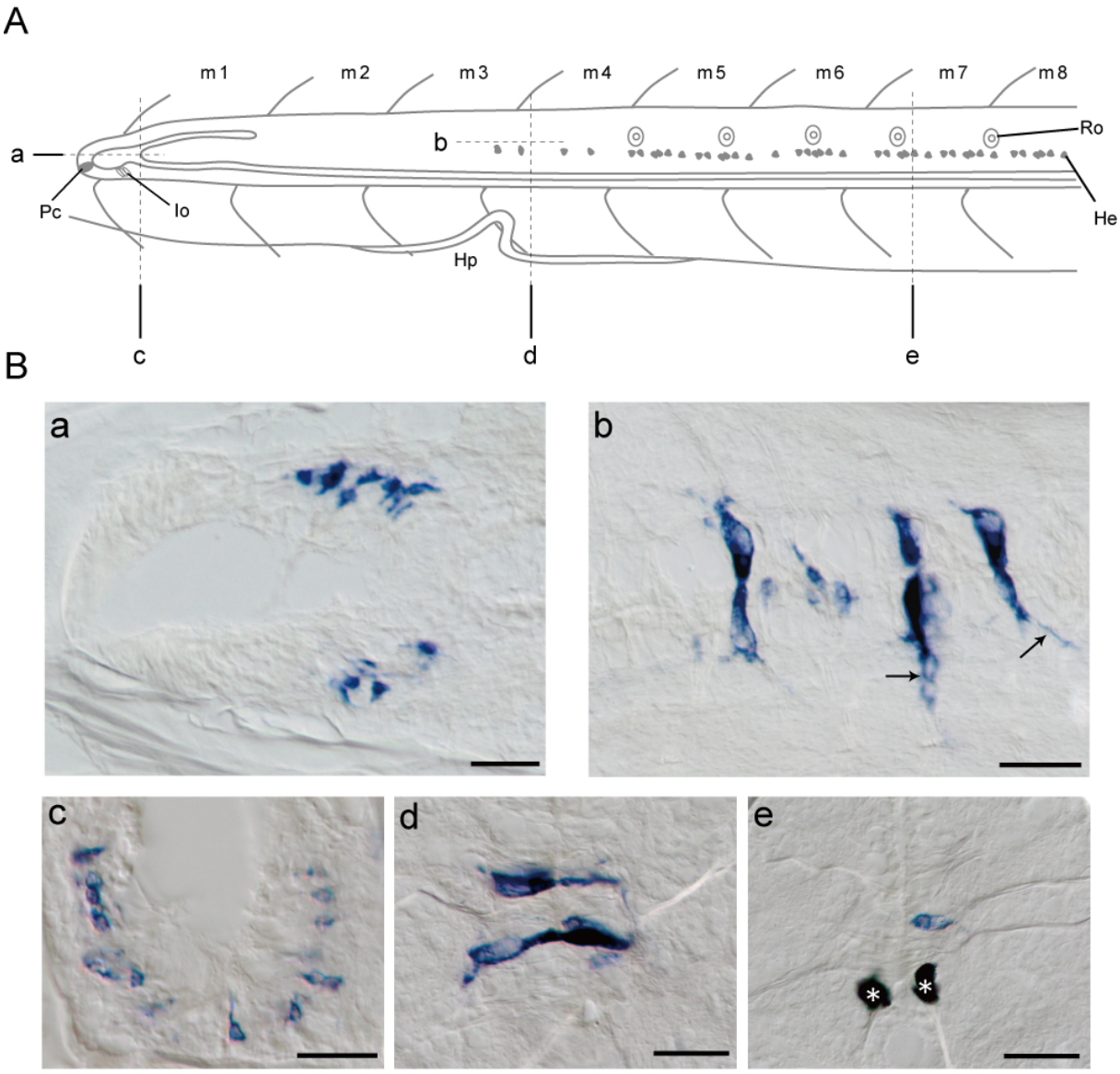

Figure 6-3. Distribution of ampVT mRNA in the anterior nerve cord of amphioxus. (A) Schematic lateral view of the anterior nerve cord showing the planes of horizontal (a, b) and transverse (c-e) frozen sections. (B) ISH signals for expression of *ampVT*. In the cerebral part of apical nerve cord, ISH positive bipolar cells that bilaterally locate around the ventral portion of central canal (a, c), while large positive cells lie across the central canal near Hatschek's pit (b, d) and the apical part of spinal cord (e). Neuronal processes run laterad in the neuropil region (arrows in b). White asterisks indicate the Hesse organs (e). He, Hesse organ; Hp, Hatschek's pit; Io, infundibular organ; m, myomere; Pc, pigment cells of frontal eye, Ro, cell of Rohde. Scale bars; 20 μm.

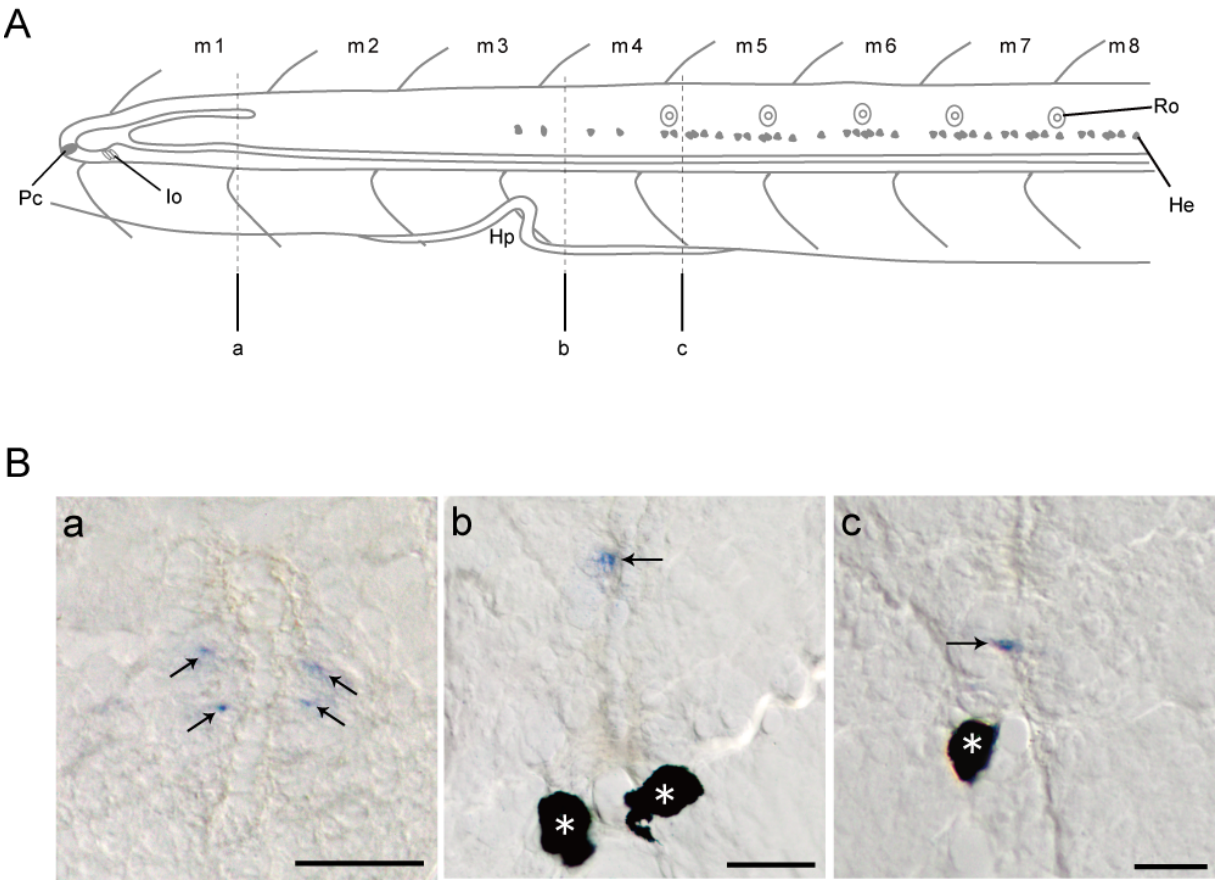

Figure 6-4. Distribution of ampGnRHR1 mRNA in the anterior nerve cord of amphioxus. (A) Schematic lateral view of the anterior nerve cord showing the planes of transverse frozen sections (a-c) at myomeres 1, 4 and 5 (m1, m4 and m5). (B) Expression of *ampGnRHR1* in neurons along both sides of the middle portion of central canal in the cerebral part of apical nerve cord (a). In the apical part of spinal cord at the levels of m4 and m5, positive signals are observed along the dorsal (b) and middle (c) portions of the central canal. Arrows indicate positive nerve cells. White asterisks indicate pigment cells (b, c). He, Hesse organ; Hp, Hatschek's pit; Io, infundibular organ; m, myomere; Pc, pigment cells of frontal eye; Ro, cell of Rohde. Scale bars, 20 μm.

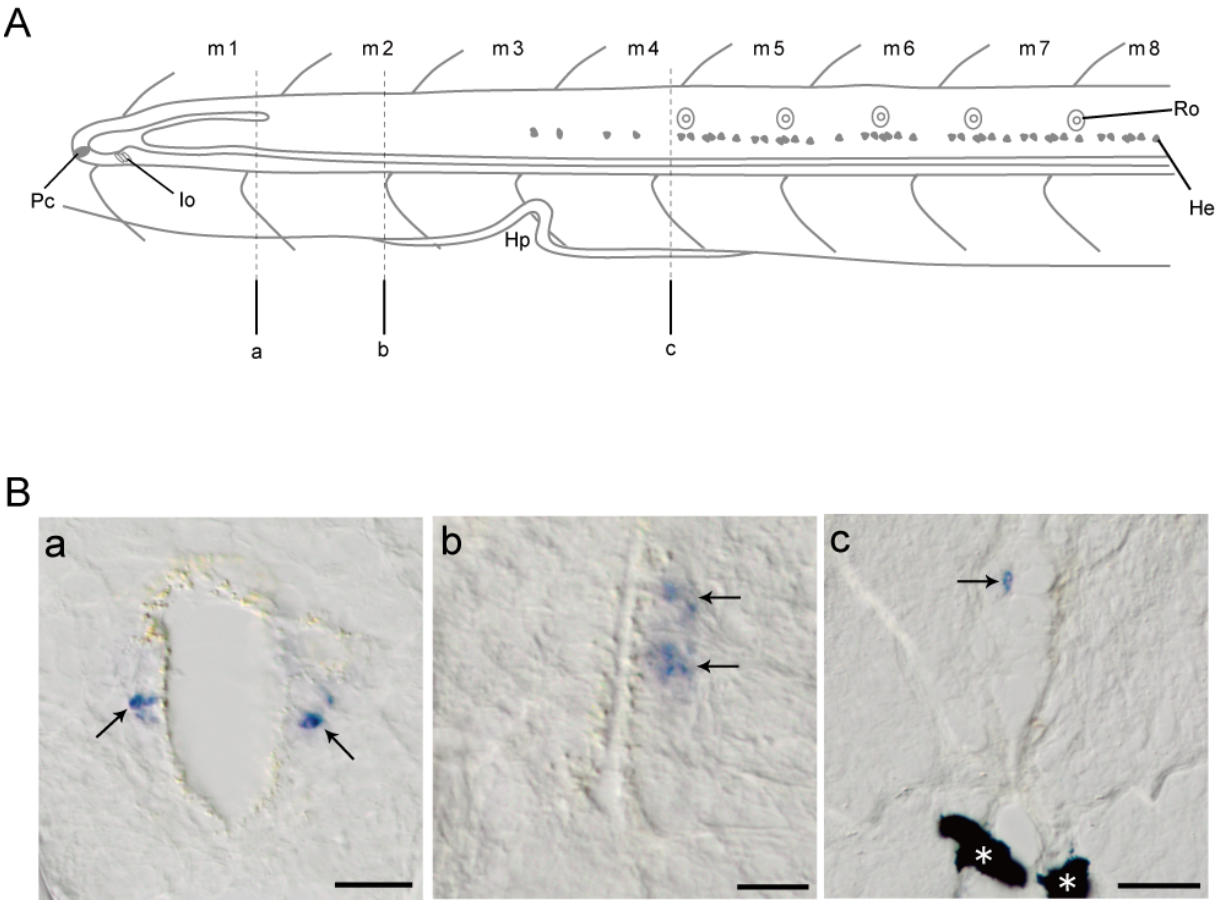

Figure 6-5. Distribution of ampGnRHR2 mRNA in the anterior nerve cord of amphioxus. (A) Schematic lateral view of the anterior nerve cord showing the planes of transverse frozen sections (a-c) at myomeres 1, 2 and 5 (m1, m2 and m5). (B) Expression of ampGnRHR2 in neurons along both sides of the middle portion of central canal in the cerebral part of apical nerve cord (a). In the regions at m2 and m5, positive cells locate along the middle and dorsal portions of the central canal (b, c). Arrows indicate positive nerve cells. White asterisks indicate Hesse organs (c). He, Hesse organ; Hp, Hatschek's pit; Io, infundibular organ; m, myomere; Pc, pigment cells of frontal eye; Ro, cell of Rohde. Scale bars, 20 μm.

Fig. 6-6.

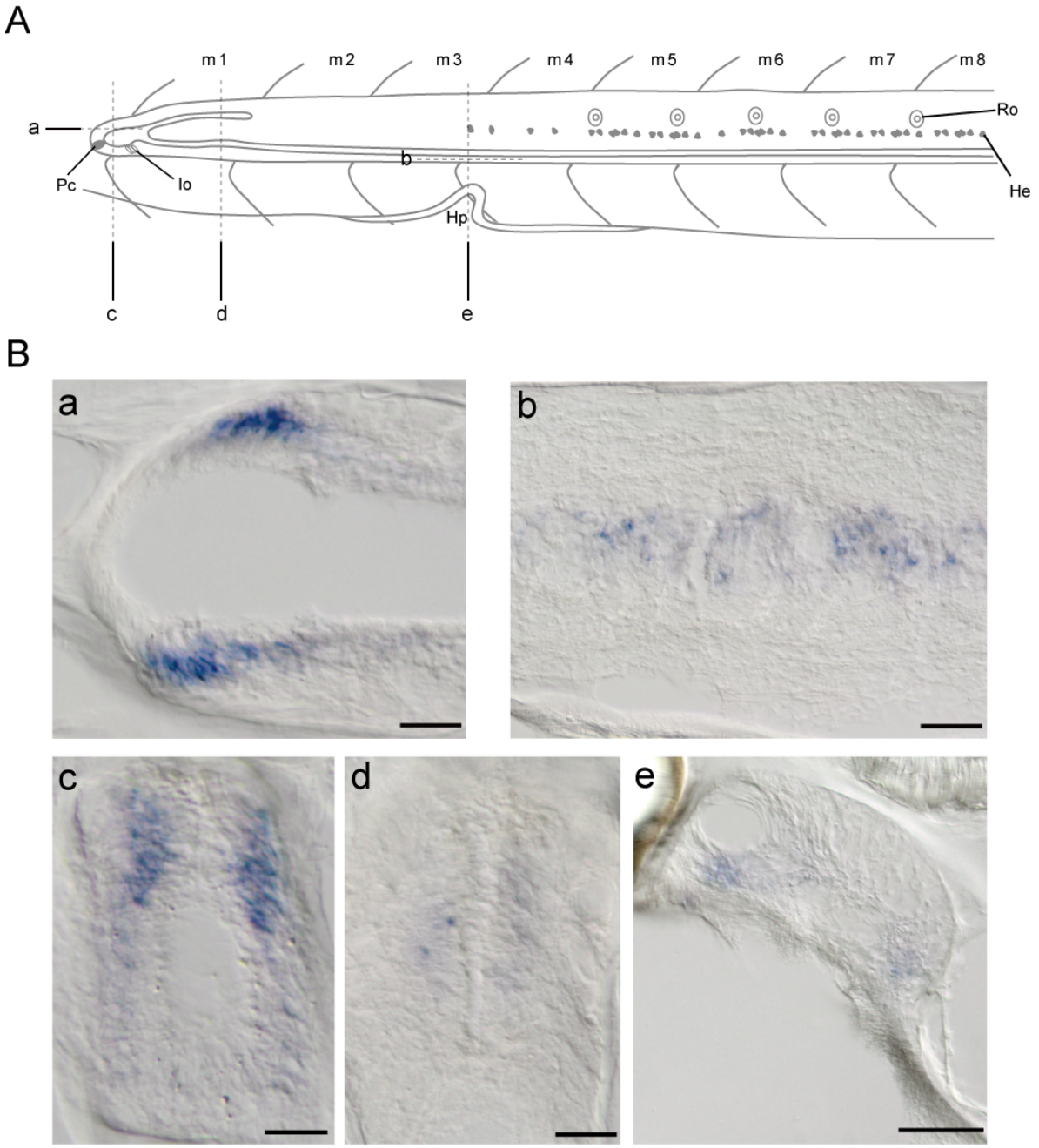

Figure 6-6. Distribution of ampER mRNA in the anterior nerve cord of amphioxus. (A) Schematic lateral view of the anterior nerve cord showing the planes of horizontal (a, b) and transverse (c-e) frozen sections in the region between the tip of cerebral part and myomere 4 (m4). (B-a, c, d) Expression of *ampER* in the cerebral part of the apical nerve cord. Note heavy signals along both sides of the dorsal portion of central canal (a, c). The number of positive cells gradually decreased in the more posterior region (d). (B-b) More caudally, a cluster of positive cells appears in the ventral portion of central canal. (B-e) Hatschek's pit bilaterally shows weak signals in the peripheral area (e). He, Hesse organ; Hp, Hatschek's pit; Io, infundibular organ; m, myomere; Pc, pigment cells of frontal eye; Ro; cell of Rohde. Scale bars, 20 μm.

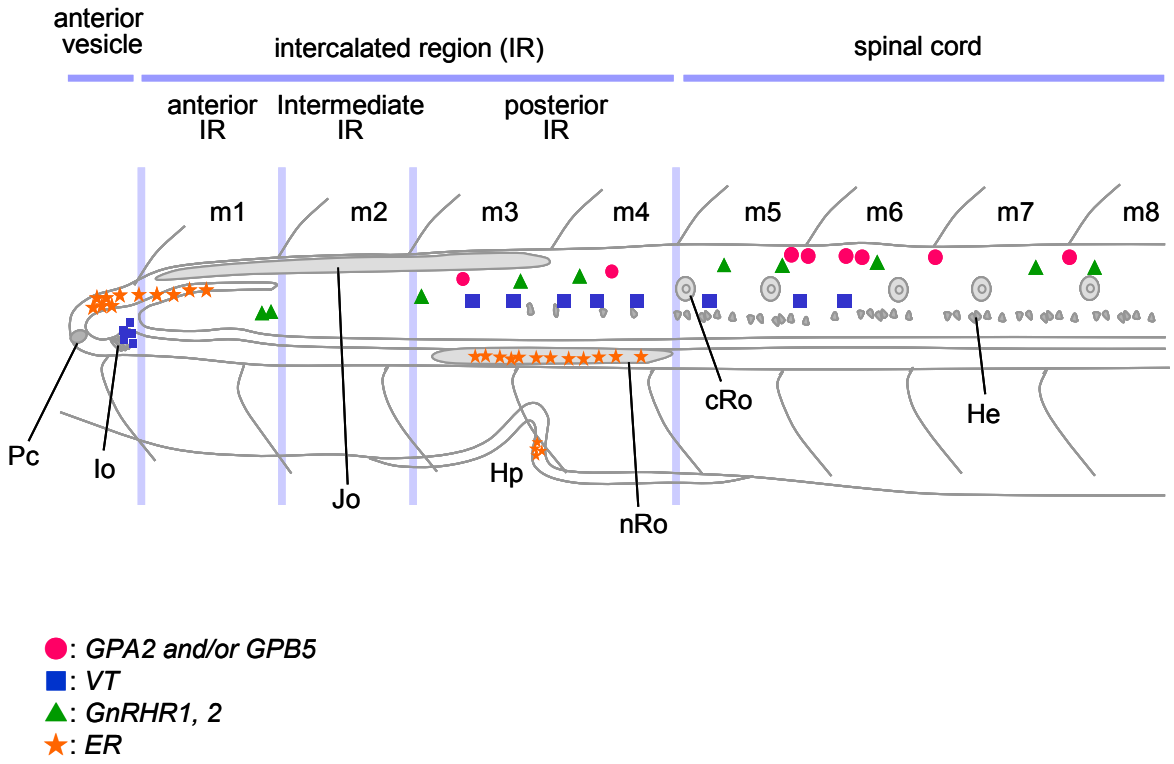

Figure 6-7. Summary diagram showing distribution of mRNAs for GPA2, GPB5, VT, GnRHR1, GnRHR2 and ER in the anterior nerve cord of amphioxus. The anterior nerve cord is divided into three regions: anterior vesicle, intercalated region (IR), and spinal cord along the rostrocaudal axis. The IR is further divided into anterior, intermediate and posterior parts. The anterior vesicle expresses *ER* and *VT*. The anterior IR expresses *ER*, *GnRHR1* and *GnRHR2*. The dorsal area of posterior IR expresses *GnRHR1*, *GnRHR2*, *GPA2* and *GPB*; the middle area, *VT*; and the ventral area, *ER*. In addition, Hatschek's pit, which lies at the junction of myomeres 3 and 4, expresses *ER*. The spinal cord expressed *GnRHR1*, *GnRHR2*, *GPA2* and *GPB5* in the dorsal area, whereas *VT* in the middle area. cRo, cell of Rhode; He, Hesse organ; Hp, Hatschek's pit; Io, infundibular organ; Jo, Joseph cells; m, myomere; nRo, nucleus of Rhode; Pc, pigment cells of frontal eye.

Chapter 7

General Discussion

The hypothalamus-pituitary axis (HPA) of vertebrates plays important roles in various physiological functions, such as regulations of metabolism, reproduction, osmoregulation, and response to stress. An abundance of studies on the HPA reported various physiological and endocrine phenomena particularly in biomedical and fisheries sciences. The previous studies on the HPA indicated that the functions of the axis are conserved among vertebrates. Even Agnatha, which is the group of the most primitive vertebrates including lamprey and hagfish, have the HPA. However, only one GTHβ in lamprey (Sower et al., 2006) and one set of gonadotropin subunits in hagfish (Nozaki et al., 2006) were reported in their pituitary glands, instead of at least four or three distinguished gonadotropin subunits in the Gnathostoma. Generally, the increase in the number of proteinaceous hormones was derived from the gene duplications during the evolution. In the case of pituitary GPHs, this event probably occurred during the evolution from the Agnatha to the Gnathostoma. This scenario leads the possibility that ancestral GPH subunit was a single molecule.

The Urochordata is the phylogenetically closest Invertebrata to the Vertebrata, and belongs to the Chordata, which is composed of the Cephalochordata, the Urochordata and the Vertebrata. The endocrine system of the Urochordata is not identical with those of the Vertebrata as well as those of other Invertebrata. However, the recent genome analyses of amphioxus *Branchiostoma floridae* in the Cephalochordata showed that 90% of amphioxus functional genes are homologous to human genes, and the gene orders on amphioxus genome scaffolds are highly similar to those on human chromosomes. Since the Cephalochordata and the Vertebrata are suggested to have evolved from a common ancestor, the survey of amphioxus endocrine system would be informative in the study of the origins of endocrine molecules and endocrine functions in the Vertebrata. However, amphioxus lacks endocrine organs such as the brain, pituitary gland, adrenal gland, spleen and thyroid gland. The analyses of amphioxus genome also showed lacks of various hormones synthesized in the

endocrine organs of the Vertebrata.

The purpose of my study was to understand the fundamental and primitive roles of the HPA in the Vertebrata through studies on amphioxus. Findings obtained from the studies on amphioxus will enrich the knowledge of the HPA of the Vertebrata. In my thesis, I surveyed the presence of amphioxus endocrine molecules on the basis of information on the HPA of the Vertebrata, and examined their actual presences in the amphioxus tissues including the nerve cord, and finally compared them with those in the Vertebrata. This approach provided several new insights into the hypothesis of evolution of the vertebrate HPA.

7-1 Evolution of pituitary glycoprotein hormones (GPHs)

A pituitary GPH is a non-covalent heterodimer with the association of two glycoprotein subunits. Each subunit forms the cystine knot structure, and the combination of two subunits construct GTHs (FSH and LH) and TSH in the Gnathostoma (Isaacs, 1995; Hearn and Gomme, 2000). The α subunit termed GPA1 is common among all GPHs, whereas homologous β subunits are specific to each hormone. Recently, a new GPH was discovered and termed thyrostimulin. This hormone consists of α subunit named as GPA2 and β subunit referred to as GPB5 (Nakabayashi et al., 2002). Two subunits, GPA2 and GPB5, exist not only in the Vertebrata, but also in the Invertebrata (Hsu et al., 2002; Park et al., 2005; Sudo et al., 2005; Dos Santos et al., 2009). In lamprey, one gonadotropin β subunit (GTHβ) was cloned and identified, and genes for GPA2 and GPB5 were found in the genome of *Petromyzon marinus* (Sower et al., 2006; 2009). It means that the Agnatha has one GTH and one thyrostimulin, while there are two GTHs and one thyrostimulin in other vertebrate groups. Accordingly, the genes for ancestral GPH are considered to be duplicated during the evolution from the Invertebrata to the Vertebrata. The following gene duplication occurred in the primitive Vertebrata, such as Agnatha, and produced two genes for α subunits (GPA1 and

GPA2) and four genes for β subunits (FSHβ, LHβ, TSHβ and GPB5) (see Fig 4-1).

The survey on nucleotide sequences in the genome database of *B. floridae* resulted in the lacks of homologous genes for the pituitary GPHs, but confirmed the presence of genes encoding GPA2 and GPB5 (Holland et al., 2008). In my study, I first reconfirmed that the amphioxus genes for GPA2 and GPB5 actually existed; however, other genes for GPHs were not found. Second, amphioxus genes for GPA2 and GPB5 were detected by a motif survey program, and then cDNAs were obtained from amphioxus, *B. belcheri*. Third, amphioxus has three GPH genes, *ampGPA2*, *ampGPA2LP*, and *ampGPB5*. Forth, in the amphioxus genome, *ampGPA2* and *ampGPB5* genes were adjacently located on the same scaffold (Dos Santos et al, 2008; Chapter 3). Furthermore, syntenies in the vicinities of GPH subunit genes were conserved between amphioxus (*B. floridae*) and human. These results suggest that *ampGPA2* and *ampGPB5* are ancestral genes for human GPH subunit genes which were probably produced by two large-scale genome duplications (Ohno, 1970), although it is suggested that the α subunit gene was lost once or twice during two genome duplications.

The chemical cross-linking experiments of recombinant ampGPA2, ampGPA2LP and ampGPB5 showed that they are able to form homodimers and heterodimers. Co-localization of ampGPA2 and ampGPB5 mRNAs was observed in the same nerve cells, and also these mRNAs were separately expressed in different nerve cells. These results suggest that amphioxus GPH is composed of two molecules with the structures of homodimer and heterodimer, and plays several roles.

As shown in the case of amphioxus GPA2 and GPB5, the increase in the number of hormone types is theoretically probable by the various combinations of a few subunits. The hypothesized idea of the evolution of pituitary GPHs in this study is summarized in Fig. 7-1. In vertebrates, the heterodimer of α and β glycoprotein subunits is known as a major form of pituitary GPHs, but a few reports on hCGβ homodimer and GPA1 homodimer (Butler et al.,

1999; Krause et al., 2007) imply the possibility of the presence of a functional homodimer in vertebrates.

7-2 Distribution of hypothalamus-pituitary axis related genes in amphioxus

The recent database on the genomic information of *B. floridae* provided us the ability of the survey of functional genes and proteins in amphioxus. Consequently, the presence or absence of homologous hormone genes related to the HPA was clarified in amphioxus (see Fig. 6-1). I obtained the candidate genes for the hormones and hormone receptors expressed in the pituitary in the amphioxus genome, and investigated the localization of transcripts of these hormone genes in the tissues and organs of amphioxus. The expressions of examined amphioxus genes were detected in the anterior part of the nerve cord of adult amphioxus. The distribution could be divided into four regions (Fig. 7-2). The first region is the anterior vesicle between the tip of the nerve cord anterior to the myomere 1, the second region is the anterior intercalated region located at the myomere 1, the third region is the posterior intercalated region between myomere 3 and 4, and the fourth is the anterior spinal cord between myomere 5 and 8.

As reported by many researchers, the nervous system of amphioxus is somewhat different from that of vertebrate. Nevertheless, Bone (1959) described the structure of nervous system in larvae, and indicated that amphioxus and vertebrates share the basic organizational features. The reports of gene expression in developmental stages of amphioxus also showed that the regulation of neural differentiation is under similar mechanisms to that seen in vertebrates, although amphioxus lacks several features of vertebrates such as formation of the telencephalon, the midbrain-hindbrain boundary region, and the neural crest (Holland and Holland, 1999; Holland and Short, 2008). Lacalli's group showed detailed anatomical characteristics of the anterior nerve cord of larva, and compared it

to the vertebrate brain in combination with the anatomical and molecular data (Wicht and Lacalli, 2005; Lacalli, 2008). However, the information of the structure and cell types in the nerve cord is quite limited in adult amphioxus. When the information on expression of genes encoding hormonal substances and receptors in amphioxus is compared with that in the hypothalamus and pituitary gland, our understanding of the characteristics of the adult nerve cord would be improved. At this point, not developmental gene expression pattern, but distribution of functional gene transcripts in the nervous system of adult animals is important for comparative study between amphioxus and vertebrates.

In my studies, expressions of the genes in the hypothalamus and the pituitary were distributed in the anterior nerve cord. This result indicates that the anterior nerve cord is the neuroendocrine center of amphioxus. Furthermore, distribution of the cell clusters which express each gene can be divided into four regions. This indicates that these discrete regions may have different functions. In particular, expression of genes found in the anterior nerve cord are distributed in the third region; the posterior intercalated region, suggest that this region in the anterior nerve cord have crucial roles for the neuroendocrine system in amphioxus.

7-3 Evolution of the hypothalamus-pituitary axis

In this study, I revealed that the anterior nerve cord is the neuroendocrine center in amphioxus. Considering this finding, I hypothesize about the evolution of the hypothalamus-pituitary endocrine axis in chordates (Fig. 7-3). The anterior nerve cord has a role for the neuroendocrine center in amphioxus. Thyrostimulin may be carried by axonal transport within the neurosecretory cells and secreted to the targets for regulation of physiological functions. During the evolution of chordates, the number of GPH increased by two round genome duplications. Furthermore, it is the key innovation, a morphological

structure of the pituitary emerged in the lineage of Vertebrata. Subsequently, the function of the secretion of GPH was shifted from the neurosecretory cells to the secretory cells in the pituitary. The present hypothalamus-pituitary axis of vertebrates has organized by the series of these exchanges.

Many text books of endocrinology tell that the endocrine system evolved from the neurosecretory system, because invertebrates have the highly developed and diverged neurosecretory systems in each animal group. Recently, supporting evidence of the evolution of the endocrine system was proposed by the molecular biological investigations of protostomes such as insects (Hartenstein et al., 2006) and annelid (Tessmar-Raible et al., 2007). These reports provided further implication for the evolution of the HPA. Furthermore, my investigation on the neuroendocrine cells in the amphioxus nerve cord supports the evolution of the endocrine system in vertebrate endocrine organs from the neurosecretory system of ancestral chordate.

The next question arising here is how endocrine substances act through the nerve cord in amphioxus. Several findings for amphioxus endocrine systems are reported. The homologues of sex steroid metabolizing enzymes and the synthetic pathway of sex steroids present in amphioxus (Mizuta and Kubokawa, 2007; Mizuta et al. 2008). Furthermore, the levels of sex steroids in the amphioxus gonads increased during the breeding season (Mizuta et al. 2008). Paris et al. (2008a; 2008b) reported that thyroid hormones of vertebrates induced the metamorphosis of amphioxus larva, and discussed the presence of a similar thyroid hormone metabolic pathway in amphioxus, although the amphioxus genome lacks the genes for thyroid hormone synthesis such as thyroglobulin. Because both of sex steroidgenesis and elevation of the thyroid hormone levels in a body are regulated by pituitary GPHs in vertebrates, these findings suggest the presence of similar endocrine regulation mechanisms of pituitary GPHs for gonads and thyroid in amphioxus. Amphioxus has one

homologous gene for GPH receptor (Dos Santos et al., 2009). The examination for the distribution of GPH receptor and physiological interaction with amphioxus thyrostimulin as well as projection of neurosecretory cells containing thyrostimulin is needed. Also, to reveal whether amphioxus has two-step endocrine regulation system such as GnRH-gonadotropin-sex steroids, identification of GnRH and TRHR which are not found yet in the genome of *B. floridae* and investigation of physiological interaction of ligand and receptor would be expected.

The emergence of the pituitary gland is the clue of the origin of the HPA. Hatschek's pit was considered to be a homologue of the pituitary gland (Hatschek, 1884; Sahlin and Olsson, 1986). However, I did not obtain any substantial evidence of the pituitary hormones in Hatschek's pit in this study. In frogs, instructive influence of the ventral diencephalon was demonstrated during the development of the pituitary primordium. When the hypothalamic infundibular primordium was removed, the presumptive pituitary develops without connection to the ventral diencephalon, but no differentiated pituitary cell types are observed (Kawamura and Kikuyama, 1998). If Hatschek's pit has the same developmental mechanism, there is a possibility that the interaction with the nerve cord and the pit are somewhat interrupted. Consequently, Hatschek's pit has a different role from that of the pituitary gland, while it may be related to secretion. I hypothesize that Hatschek's pit is the primordial pituitary, but that the pit does not acquire the functions of the pituitary. Further investigation of pituitary development, specifically first interaction with the diencephalon, might provide the evolutional origin of the pituitary.

In conclusion, my study provided an insight of how the vertebrate HPA evolved. The evolutionary investigation on the amphioxus endocrine system improves the fundamental knowledge of the endocrine system in vertebrates. Needless to say, progress of applied sciences is not possible without basic sciences. At this point, my research will contribute to

understanding of developmental and functional mechanisms of the pituitary gland, as well as the application for medical research related to diseases of the hypothalamus and pituitary, and for fisheries research on the culture of aquatic organisms and control of the fishery resources by focusing on reproduction, development, growth and adaptation.

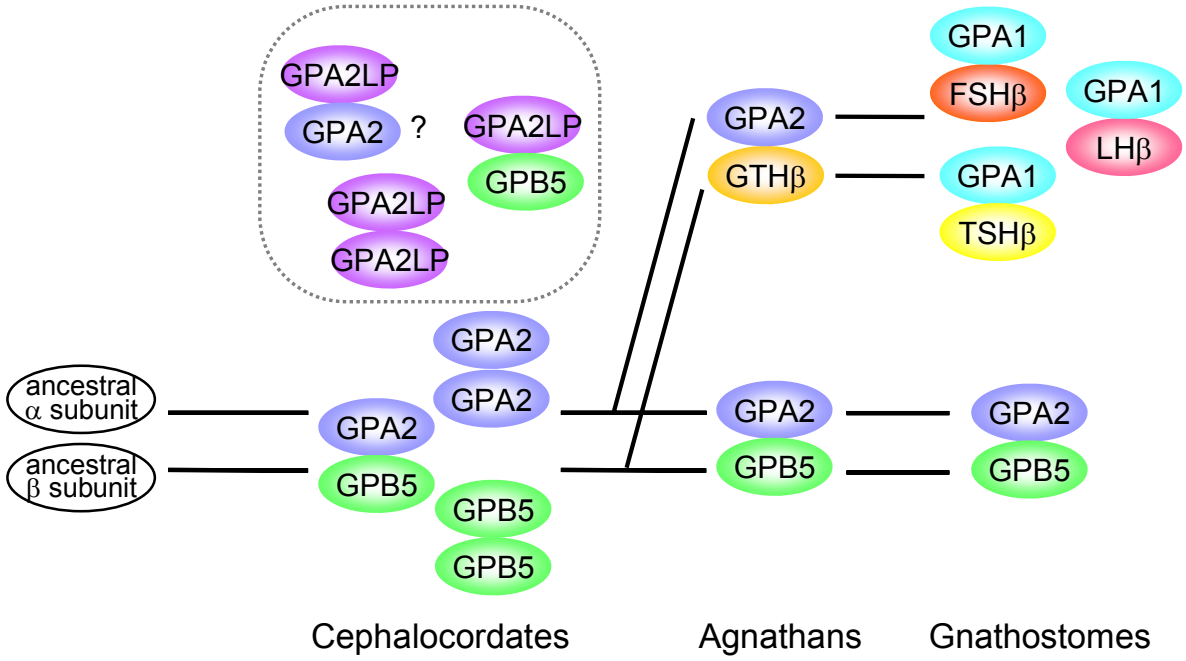

Figure 7-1. Scheme for evolution of glycoprotein hormones (GPH). Cephalochordates have GPA2 as α subunit and GPB5 as β subunit. They can form both homo- and hetero-dimers. GPA2LP also may act as a homo or a heheto-dimer. My present study indicates that the pituitary GPH subunits first arose in agnathans from ancestral GPH subunits, GPA2 and GPB5. Gene duplications after the agnathan-gnathostomes divergence produced GPA1 and three GPBs.

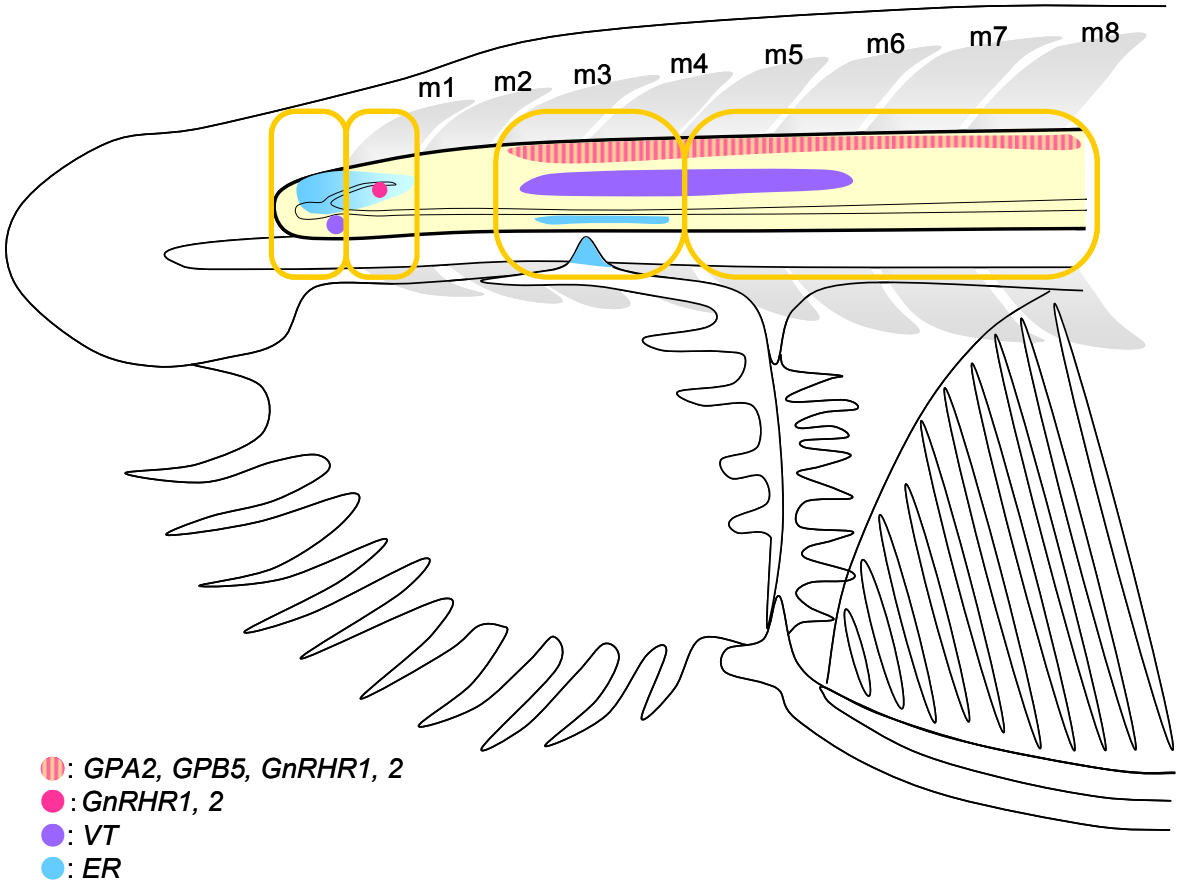

Figure 7-2. Distributions of the gene transcripts investigated in this study. The anterior nerve cord may function as a neuroendocrine center in amphioxus. Discrete four cell groups described in Fig 6-7 may have different roles.

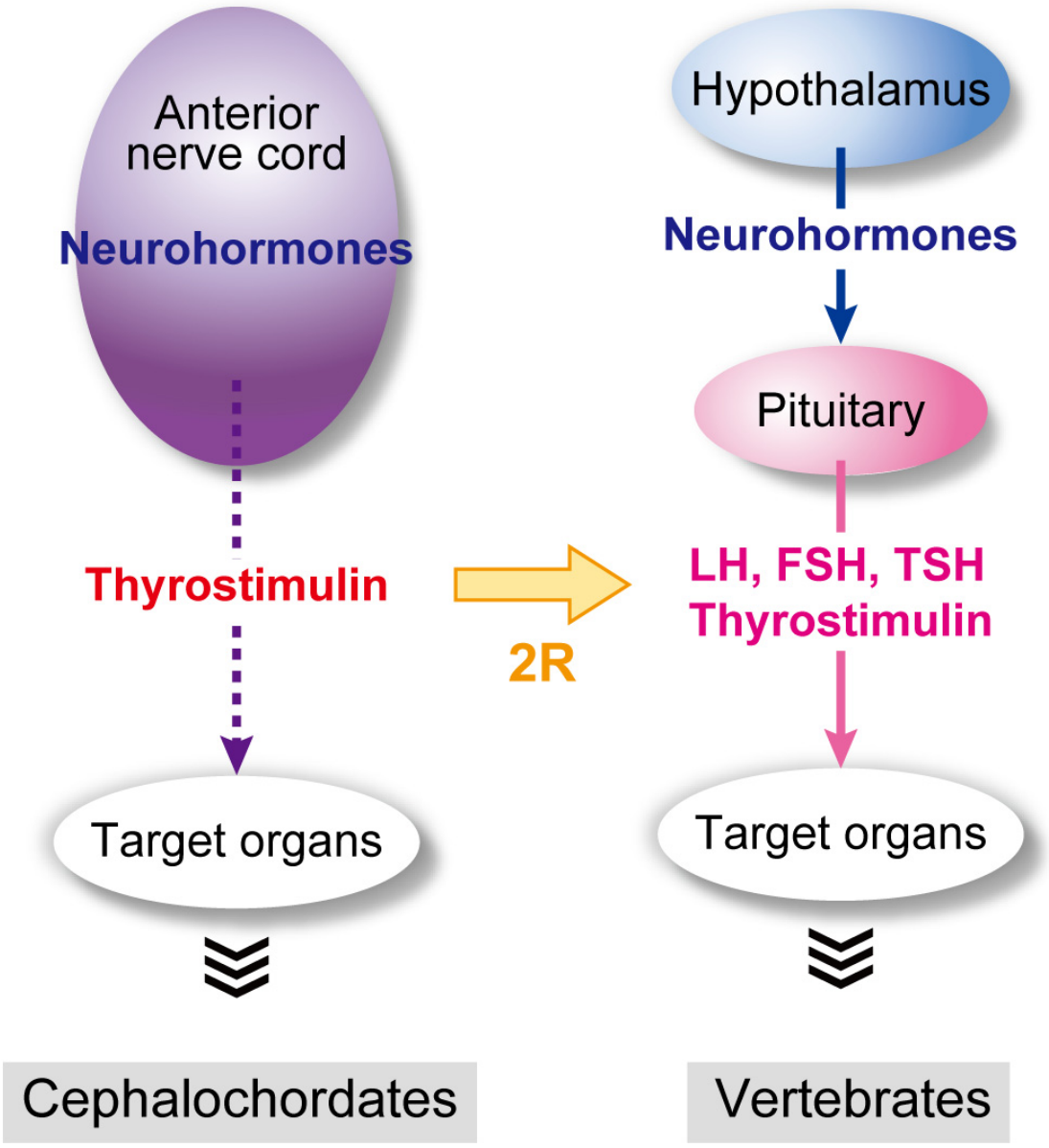

Figure 7-3. Comparison of the neuroendocrine systems between cephalochordates and vertebrates. Cephalochordates develop the neuroendocrine system in the anterior nerve cord, which secretes thyrostimulin from nerve terminals of nerve cells in the anterior nerve cord, and regulate physiological functions of peripheral target organs, such as gonads. Although evolutionary process is not known yet, this system is considered to yield the increased number of glycoprotein hormones by two rounds of genome duplication in accordance with the emergence of an ancestral pituitary gland in primitive vertebrates.

References

Anadon R, Adrio F, Rodriguez-Moldes I. 1998. Distribution of GABA immunoreactivity in the central and peripheral nervous system of amphioxus (*Branchiostoma lanceolatum* Pallas). J Comp Neurol 401(3):293-307.

Barrington EJW. 1963. An Introduction to general and comparative endocrinology. Clarendon Press, Oxford.

Begeot M, Hemming FJ, Dubois PM, Combarnous Y, Dubois MP, Aubert ML. 1984. Induction of pituitary lactotrope differentiation by luteinizing hormone alpha subunit. Science 226 (4674):566-568.

Blair JE, Hedges SB. 2005. Molecular phylogeny and divergence times of deuterostome animals. Mol Biol Evol 22(11):2275-2284.

Bollenbacher WE, Vedeckis WV, Gilbert LI. 1975. Ecdysone titers and prothoracic gland activity during the larval-pupal development of *Manduca sexta*. Dev Biol 44(1):46-53.

Bone Q. 1959. The central nervous system in larval acraniates. Quar J Micr Sci 100:509-527.

Bone Q. 1960. The central nervous system in amphioxus. J Comp Neurol 115:27-64.

Boorman CJ, Shimeld SM. 2002. Pitx homeobox genes in Ciona and amphioxus show left-right asymmetry is a conserved chordate character and define the ascidian adenohypophysis. Evol Dev 4(5):354-365.

Bridgham JT, Brown JE, Rodr guez-Mar A, Catchen JM, Thornton JW. 2008. Evolution of a new function by degenerative mutation in cephalochordate steroid receptors. PLoS Genet 4(9):e1000191.

Butler SA, Laidler P, Porter JR, Kicman AT, Chard T, Cowan DA, Iles RK. 1999. The β-subunit of human chorionic gonadotrophin exists as a homodimer. J Mol Endocrinol 22(2):185-192.

Candiani S, Holland ND, Oliveri D, Parodi M, Pestarino M. 2008. Expression of the amphioxus Pit-1 gene (AmphiPOU1F1/Pit-1) exclusively in the developing preoral organ, a putative homolog of the vertebrate adenohypophysis. Brain Res Bull 75(2-4):324-330.

Candiani S, Pestarino M. 1998. Expression of the tissue-specific transcription factor Pit-1 in the lancelet, *Branchiostoma lanceolatum*. J Comp Neurol 392(3):343-351.

Castro A, Manso MJ, Anadon R. 2003. Distribution of neuropeptide Y immunoreactivity in the central and peripheral nervous systems of amphioxus (*Branchiostoma lanceolatum* Pallas). J Comp Neurol 461(3):350-361.

Chang CY, Chu YT, Chen D. 1982. Immunocytochemical demonstration of luteinizing hormone (LH) in Hatschek's pit of amphioxus *(Branchistoma belcheri)*. Kexue Tongbao 27:1233-1234.

Chang CY, Liu YH, Zhu YT, Zhu HH. 1985. The reproductive endocrinology of Amphioxus. In: Carlick DG, Korner PI, eds Frontiers in physiological research Canberra, Australia: Australian Academy of Science:79–86.

Civelli O, Birnberg N, Herbert E. 1982. Detection and quantitation of pro-opiomelanocortin mRNA in pituitary and brain tissues from different species. J Biol Chem 257(12):6783-6787.

Croxatto H, Arrau J, Croxatto H. 1964. Luteinizing hormone-like activity in human median eminence extracts. Nature 204(4958):584-585.

Cruz LJ, de Santos V, Zafaralla GC, Ramilo CA, Zeikus R, Gray WR, Olivera BM. 1987. Invertebrate vasopressin/oxytocin homologs. Characterization of peptides from *Conus geographus* and *Conus straitus venoms*. J Biol Chem 262(33):15821-15824.

Delsuc F, Brinkmann H, Chourrout D, Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature 439(7079):965-968.

Dos Santos S, Bardet C, Bertrand S, Escriva H, Habert D, Querat B. 2009. Distinct expression patterns of glycoprotein hormone-α2 and -β5 in a basal chordate suggest independent developmental functions. Endocrinology 150(8):3815-3822.

Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD, Sorensen MV, Haddock SH, Schmidt-Rhaesa A, Okusu A, Kristensen RM, Wheeler WC, Martindale MQ, Giribet G. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452(7188):745-749.

Ekhart D, Korf HW, Wicht H. 2003. Cytoarchitecture, topography, and descending supraspinal projections in the anterior central nervous system of *Branchiostoma lanceolatum*. J Comp Neurol 466(3):319-330.

Emanuele N, Connick E, Howell T, Anderson J, Hojvat S, Baker G, Souchek J, Kirsteins L, Lawrence AM. 1981. Hypothalamic luteinizing hormone (LH): characteristics and response to hypophysectomy. Biol Reprod 25(2):321-326.

Fan QR, Hendrickson WA. 2005. Structure of human follicle-stimulating hormone in complex with its receptor. Nature 433(7023):269-277.

Fang YQ, Wang L. 1984. The preliminary study of the wheel organ and Hatschek's pit of amphioxus on testicular development in young toad (*Bufo melanostictus*). Acta Biol Exp Sinica 17:115-117.

Franz V. 1923. Haut, Sinnesorgane und Nervensystem der Akranier. Jena Z Naturw 59:402–526.

Franz V. 1927. Morpohologie der Akranier. Ergebn Anat Entwgech 27:465-692.

Freeman ME, Kanyicska B, Lerant A, Nagy G. 2000. Prolactin: structure, function, and regulation of secretion. Phys Rev 80(4):1523-1631.

Garcia-Fernandez J, Holland PW. 1994. Archetypal organization of the amphioxus Hox gene cluster. Nature 370(6490):563-566.

Garcia-Fernandez J, Benito-Gutierrez E. 2009. It's a long way from amphioxus: descendants of the earliest chordate. Bioessays 31(6):665-675.

Glardon S, Holland LZ, Gehring WJ, Holland ND. 1998. Isolation and developmental expression of the amphioxus Pax-6 gene (AmphiPax-6): insights into eye and photoreceptor evolution. Development 125(14):2701-2710.

Goodrich ES. 1918. "Proboscis pores" in craniate vertebrates, a auggenstion concerning the premandibular somites and hypophysis. Quart J Microsc 62:539-553.

Gorbman A. 1965. Vascular relations between the neurohypophysis and adenohypophysis of cyclostomes and the problem of evolution of hypothalamic neuroendrocrine control.

Arch Anat Microsc Morphol Exp 54:163-194.
Gorbman A, Dickhoff WW, Vigna SR, Clark NB, Ralph CL. 1983. Comparative endocrinology. John Wiley & Sons, New York.
Gorbman A, Nozaki M, Kubokawa K. 1999. A brain-Hatschek's pit connection in amphioxus. Gen Comp Endocrinol 113(2):251-254.
Green JD. 1951. The comparative anatomy of the hypophysis with special reference to its blood suplly and innervation. Amer J Anatom 88(2):225-311.
Gwee PC, Tay BH, 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.
Hartenstein V. 2006. The neuroendocrine system of invertebrates: a developmental and evolutionary perspective. J Endocrinol 190(3):555-570.
Hatschek B. 1884. Mitteilungen über Amphioxus. Zool Anz 7:86-127.
Hauser F, Williamson M, Cazzamali G, Grimmelikhuijzen CJ. 2006. Identifying neuropeptide and protein hormone receptors in *Drosophila melanogaster* by exploiting genomic data. Brief. Funct. Genomic Proteomics 4(4):321-330.
Hawkins MB, Thornton JW, Crews D, Skipper JK, Dotte A, Thomas P. 2000. Identification of a third distinct estrogen receptor and reclassification of estrogen receptors in teleosts. Proc Natl Acad Sci 97(20):10751-10756.
Hearn MT, Gomme PT. 2000. Molecular architecture and biorecognition processes of the cystine knot protein superfamily: part I. The glycoprotein hormones. J Mol Recognit 13(5):223-278.
Hojvat S, Baker G, Kirsteins L, Lawrence AM. 1982. TSH in the rat and monkey brain. Distribution, characterization and effect of hypophysectomy. Neuroendocrinology 34(5):327-332.
Holland LZ, Albalat R, Azumi K, Benito-Gutierrez E, Blow MJ, Bronner-Fraser M, Brunet F, Butts T, Candiani S, Dishaw LJ, Ferrier DE, Garcia-Fernandez J, Gibson-Brown JJ, Gissi C, Godzik A, Hallbook F, Hirose D, Hosomichi K, Ikuta T, Inoko H, Kasahara M, Kasamatsu J, Kawashima T, Kimura A, Kobayashi M, Kozmik Z, Kubokawa K, Laudet V, Litman GW, McHardy AC, Meulemans D, Nonaka M, Olinski RP, Pancer Z, Pennacchio LA, Pestarino M, Rast JP, Rigoutsos I, Robinson-Rechavi M, Roch G, Saiga H, Sasakura Y, Satake M, Satou Y, Schubert M, Sherwood N, Shiina T, Takatori N, Tello J, Vopalensky P, Wada S, Xu A, Ye Y, Yoshida K, Yoshizaki F, Yu JK, Zhang Q, Zmasek CM, de Jong PJ, Osoegawa K, Putnam NH, Rokhsar DS, Satoh N, Holland PW. 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Res 18(7):1100-1111.
Holland LZ, Holland ND. 1999. Chordate origins of the vertebrate central nervous system. Curr Opin Neurobiol 9(5):596-602.
Holland LZ, Short S. 2008. Gene duplication, co-option and recruitment during the origin of the vertebrate brain from the invertebrate chordate brain. Brain Behav Evol 72(2):91-105.

Hostetter G, Eaton A, Carnes M, Gildner J, Brownfield MS. 1987. Immunocytochemical distribution of luteinizing hormone in rat central nervous system. Neuroendocrinology 46(3):185-193.

Hsu SY, Nakabayashi K, Bhalla A. 2002. Evolution of glycoprotein hormone subunit genes in bilateral metazoa: identification of two novel human glycoprotein hormone subunit family genes, GPA2 and GPB5. Mol Endocrinol 16(7):1538-1551.

Iles RK. 2007. Ectopic hCGβ expression by epithelial cancer: malignant behaviour, metastasis and inhibition of tumor cell apoptosis. Mol Cell Endocrinol 260-262:264-270.

Isaacs NW. 1995. Cystine knots. Curr Opin Struct Biol 5(3):391-395.

Jacobson AG, Miyamoto DM, Mai SH. 1979. Rathke's pouch morphogenesis in the chick embryo. J Exp Zool 207:351-366.

Kagawa H, Kawazoe I, Tanaka H, Okuzawa K. 1998. Immunocytochemical identification of two distinct gonadotropic cells (GTH I and GTH II) in the pituitary of bluefin tuna, *Thunnus thynnus*. Gen Comp Endocrinol 110(1):11-18.

Kah O, Lethimonier C, Somoza G, Guilgur LG, Vaillant C, Lareyre JJ. 2007. GnRH and GnRH receptors in metazoa: a historical, comparative, and evolutive perspective. Gen Comp Endocrinol 153(1-3):346-364.

Kanda A, Takahashi T, Satake H, Minakata H. 2006. Molecular and functional characterization of a novel gonadotropin-releasing-hormone receptor isolated from the common octopus (*Octopus vulgaris*). Biochem J 395(1):125-135.

Kato T, Kajikawa M, Maenaka K, Park EY. 2009. Silkworm expression system as a platform technology in life science. Appl Microbiol Biotechnol.

Katsu Y, Kubokawa K, Urushitani H, Iguchi T. 2009. Estrogen-dependent transactivation of amphioxus steroid hormone receptor via both estrogen and androgen response elements. Endocrinology.

Katsukura Y, Ando H, David CN, Grimmelikhuijzen CJ, Sugiyama T. 2004. Control of planula migration by LWamide and RFamide neuropeptides in *Hydractinia echinata*. J Exp Biol 207(11):1803-1810.

Kawamura K, Kikuyama S. 1998. Morphogenesis of the hypothalamus and hypophysis: their association, dissociation and reassociation before and after "Rathke". Arch Histol Cytol 61(3):189-198.

Keller R. 1992. Crustacean neuropeptides: structures, functions and comparative aspects. Experientia 48(5):439-448.

Kobayashi H, Uemura H. 1972. The neurohypophysis of the hagfish, *Eptatretus burgeri* (Girard). Gen Comp Endocrinol 3(Suppl 1):114-124.

Kotani H, Hoshimaru M, Nawa H, Nakanishi S. 1986. Structure and gene organization of bovine neuromedin K precursor. Proc Natl Acad Sci 83(18):7074-7078

Krause JM, Berger P, Roig J, Singh V, Merz WE. 2007. Rapid maturation of glycoprotein hormone free α-subunit (GPHα) and GPHαα homodimers. Mol Endocrinol 21(10):2551-2564.

Kumar S, Tamura K, Nei M. 2004. MEGA3: Integrated software for molecular evolutionary

genetics analysis and sequence alignment. Brief Bioinform 5(2):150-163.

Kubokawa K, Azuma N, Tomiyama M. 1998. A new population of the amphioxus (*Branchiostoma belcheri*) in the Enshu-Nada Sea in Japan. Zool Sci 15(5):799-803.

Kubokawa K, Mizuta T, Morisawa M, Azuma N. 2003. Gonadal state of wild amphioxus populations and spawning success in captive conditions during the breeding period in Japan. Zool Sci 20(7):889-895.

Kusakabe T, Mishima S, Shimada I, Kitajima Y, Tsuda M. 2003. Structure, expression, and cluster organization of genes encoding gonadotropin-releasing hormone receptors found in the neural complex of the ascidian *Ciona intestinalis*. Gene 322:77-84.

Lacalli TC. 1996. Frontal eye circuitry, rostral sensory pathways and brain organization in amphioxus larvae: evidence from 3D reconstructions. Phil Trans R Soc B 351(1337):243-263.

Lacalli TC. 2008. Basic features of the ancestral chordate brain: a protochordate perspective. Brain Res Bull 75(2-4):319-323.

Lacalli TC, Holland ND, West JE. 1994. Landmarks in the anterior central nervous system of amphioxus larvae. Phil Trans R Soc B 344(1308):165-185.

Ladanyi A, Sipos F, Szoke D, Galamb O, Molnar B, Tulassay Z. 2006. Laser microdissection in translational and clinical research. Cytometry A 69(9):947-960.

Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008. Glycosyltransferases: structures, functions, and mechanisms. Ann Rev Biochem 77:521-555.

Larsson TA, Olsson F, Sundstrom G, Lundin LG, Brenner S, Venkatesh B, Larhammar D. 2008. Early vertebrate chromosome duplications and the evolution of the neuropeptide Y receptor gene regions. BMC Evol Biol 8(1):184.

Li B, Predel R, Neupert S, Hauser F, Tanaka Y, Cazzamali G, Williamson M, Arakane Y, Verleyen P, Schoofs L, Schachtner J, Grimmelikhuijzen CJ, Park Y. 2008. Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. Genome Res 18(1):113-122.

Li MD, Ford JJ. 1998. A comprehensive evolutionary analysis based on nucleotide and amino acid sequences of the α- and β-subunits of glycoprotein hormone gene family. J Endocrinol 156(3):529-542.

Macdonald LE, Wortley KE, Gowen LC, Anderson KD, Murray JD, Poueymirou WT, Simmons MV, Barber D, Valenzuela DM, Economides AN, Wiegand SJ, Yancopoulos GD, Sleeman MW, Murphy AJ. 2005. Resistance to diet-induced obesity in mice globally overexpressing OGH/GPB5. Proc Natl Acad Sci 102(7):2496-2501.

Mackie GO, Marx RM, Meech RW. 2003. Central circuitry in the jellyfish *Aglantha digitale* IV. Pathways coordinating feeding behaviour. J Exp Biol 206(Pt 14):2487-2505.

Massari M, Candiani S, Pestarino M. 1999. Distribution and localization of immunoreactive FMRFamide-like peptides in the lancelet. Eur J Histochem 43:63-39.

Meves A. 1973. Elektronenmikroskopische untersuchungen über die zytoarchitektur des gehirns von *Branchiostoma lanceolatum*. Z Zellforsch 139:511-532.

Mizuta T, Asahina K, Suzuki M, Kubokawa K. 2008. In vitro conversion of sex steroids and

expression of sex steroidogenic enzyme genes in amphioxus ovary. Journal of experimental zoology 309(2):83-93.

Mizuta T, Kubokawa K. 2004. Non-synchronous spawning behavior in laboratory reared amphioxus *Branchiostoma belcheri* Gray. The Journal of experimental biology 309(2):239-251.

Mizuta T, Kubokawa K. 2007. Presence of sex steroids and cytochrome P450 genes in amphioxus. Endocrinology 148(8):3554-3565.

Moore RY, Speh JC. 1993. GABA is the principal neurotransmitter of the circadian system. Neurosci Lett 150(1):112-116.

Moriarty GC. 1973. Adenohypophysis: ultrastructural cytochemistry. A review. J Histochem Cytochem 21(10):855-894.

Moriarty GC. 1976. Immunocytochemistry of the pituitary glycoprotein hormones. J Histochem Cytochem 24(7):846-863.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35 (Web Server issue):W182-185.

Moshitzky P, Fleischmann I, Chaimov N, Saudan P, Klauser S, Kubli E, Applebaum SW. 1996. Sex-peptide activates juvenile hormone biosynthesis in the *Drosophila melanogaster* corpus allatum. Arch Insect Biochem Physiol 32(3-4):363-374.

Motohashi T, Shimojima T, Fukagawa T, Maenaka K, Park EY. 2005. Efficient large-scale protein production of larvae and pupae of silkworm by *Bombyx mori* nuclear polyhedrosis virus bacmid system. Biochem Biophys Res Commun 326(3):564-569.

Nagasaki H, Wang Z, Jackson VR, Lin S, Nothacker HP, Civelli O. 2006. Differential expression of the thyrostimulin subunits, glycoprotein α2 and β5 in the rat pituitary. J Mol Endocrinol 37(1):39-50.

Nakabayashi K, Matsumi H, Bhalla A, Bae J, Mosselman S, Hsu SY, Hsueh AJ. 2002. Thyrostimulin, a heterodimer of two new human glycoprotein hormone subunits, activates the thyroid-stimulating hormone receptor. J Clin Invest 109(11):1445-1452.

Nakane PK. 1970. Classifications of anterior pituitary cell types with immunoenzyme histochemistry. J Histochem Cytochem 18(1):9-20.

Nakanishi S, Inoue A, Kita T, Nakamura M, Chang AC, Cohen SN, Numa S. 1979. Nucleotide sequence of cloned cDNA for bovine corticotropin-β-lipotropin precursor. Nature 278(5703):423-427.

Nemansky M, Moy E, Lyons CD, Yu I, Blithe DL. 1998. Human endometrial stromal cells generate uncombined α-subunit from human chorionic gonadotropin, which can synergize with progesterone to induce decidualization. J Clin Endocrinol Metab 83(2):575-581.

Nozaki M, Gorbman A. 1992. The question of functional homology of Hatschek's pit of amphioxus *(Branchiostoma belcheri)* and the vertebrate adenohypophysis. Zool Sci 9:387-395.

Nozaki M, Gorbman A, Sower SA. 1994. Diffusion between the neurohypophysis and the

adenohypophysis of lampreys, *Petromyzon marinus*. Gen Comp Endocrinol 96(3):385-391.

Nozaki M, Naito N, Swanson P, Miyata K, Nakai Y, Oota Y, Suzuki K, Kawauchi H. 1990. Salmonid pituitary gonadotrophs. I. Distinct cellular distributions of two gonadotropins, GTH I and GTH II. Gen Comp Endocrinol 77(3):348-357.

Nozaki M, Terakado K, Kubokawa K. 1999. Unique asymmetric protrusion of nerve cord in the Amphioxus, *Branchiostoma belcheri*. Naturwissenschaften 86(7):328-330.

Nozaki M, Uchida K, Miki M, Shimotani T, Honda K, Moriyama S, Kawauchi H. 2006. Molecular cloning and identification of hagfish gonadotropin (GTH). Zool Sci 23(12):1211-1211.

Ogasawara M, Satoh N, Shimada Y, Wang Z, Tanaka T, Noji S. 2006. Rapid and stable buffer exchange system using InSitu Chip suitable for multicolor and large-scale whole-mount analyses. Dev Genes Evol 216(2):100-104.

Ohno S. 1970. Evolution by genge duplication. Springer-Verlag.

Okada SL, Ellsworth JL, Durnam DM, Haugen HS, Holloway JL, Kelley ML, Lewis KE, Ren H, Sheppard PO, Storey HM, Waggie KS, Wolf AC, Yao LY, Webster PJ. 2006. A glycoprotein hormone expressed in corticotrophs exhibits unique binding properties on thyroid-stimulating hormone receptor. Mol Endocrinol 20(2):414-425.

Okajima Y, Nagasaki H, Suzuki C, Suga H, Ozaki N, Arima H, Hamada Y, Civelli O, Oiso Y. 2008. Biochemical roles of the oligosaccharide chains in thyrostimulin, a heterodimeric hormone of glycoprotein hormone subunits alpha 2 (GPA2) and beta 5 (GPB5). Regul Pept 148(1-3):62-67.

Okamoto M, Tsutsui S, Tasumi S, Suetake H, Kikuchi K, Suzuki Y. 2005. Tandem repeat L-rhamnose-binding lectin from the skin mucus of ponyfish, Leiognathus nuchalis. Biochem Biophys Res Commun 333(2):463-469.

Oumi T, Ukena K, Matsushima O, Ikeda T, Fujita T, Minakata H, Nomoto K. 1994. Annetocin: an oxytocin-related peptide isolated from the earthworm, *Eisenia foetida*. Biochem Biophys Res Commun 198(1):393-399.

Pallas P. 1774. *Limax lanceolatus.* Descriptio Limacis lanceolaris. In: Spicilegia Zoologica, quibus novae imprimus et obscurae animalium species iconibus, descriptiomibus Gottl August Lange, Berlin, 10, pp 1-19.

Parhar IS, Soga T, Ogawa S, Sakuma Y. 2003. FSH and LH-b subunits in the preoptic nucleus: ontogenic expression in teleost. Gen Comp Endocrinol 132(3):369-378.

Paris M, Brunet F, Markov GV, Schubert M, Laudet V. 2008a. The amphioxus genome enlightens the evolution of the thyroid hormone signaling pathway. Dev Genes Evol 218(11-12):667-680.

Paris M, Escriva H, Schubert M, Brunet F, Brtko J, Ciesielski F, Roecklin D, Vivat-Hannah V, Jamin EL, Cravedi JP, Scanlan TS, Renaud JP, Holland ND, Laudet V. 2008b. Amphioxus postembryonic development reveals the homology of chordate metamorphosis. Curr Biol 18(11):825-830.

Paris M, Pettersson K, Schubert M, Bertrand S, Pongratz I, Escriva H, Laudet V. 2008c. An

amphioxus orthologue of the estrogen receptor that does not bind estradiol: insights into estrogen receptor evolution. BMC Evol Biol 8(1):219.

Park JI, Semyonov J, Chang CL, Hsu SY. 2005. Conservation of the heterodimeric glycoprotein hormone subunit family proteins and the LGR signaling system from nematodes to humans. Endocrine 26(3):267-276.

Pestarino M, Lucaroni B. 1996. FMRFamide-like immunoreactivity in the central nervous system of the lancelet *Branchiostoma lanceolatum*. Isr J Zool 42:227–234.

Popa G. 1930. A Portal circulation from the pituitary to the hypothalamic region. J Anat 65:88-91.

Proux JP, Miller CA, Li JP, Carney RL, Girardie A, Delaage M, Schooley DA. 1987. Identification of an arginine vasopressin-like diuretic hormone from *Locusta migratoria*. Biochem Biophys Res Commun 149(1):180-186.

Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutierrez EL, Dubchak I, Garcia-Fernandez J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov VV, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin IT, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PW, Satoh N, Rokhsar DS. 2008. The amphioxus genome and the evolution of the chordate karyotype. Nature 453(7198):1064-1071.

Quérat B, Sellouka A, Salmon C. 2000. Phylogenetic analysis of the vertebrate glycoprotein hormone family including new sequences of sturgeon (*Acipenser baeri*) β subunits of the two gonadotropins and the thyroid-stimulating hormone. Biol Reprod 63:222-228.

Quérat B, Tonnerre-Doncarli C, Genies F, Salmon C. 2001. Duality of gonadotropins in gnathostomes. Gen Comp Endocrinol 124(3):308-314.

Reich G. 1992. A new peptide of the oxytocin/vasopressin family isolated from nerves of the cephalopod *Octopus vulgaris*. Neurosci Lett 134(2):191-194.

Rose MC. 1992. Mucins: structure, function, and role in pulmonary diseases. Am J Physiol 263:L413-429.

Ruppert EE. 1997. Cephalochordata (Acrania). In: Harrison FW, Ruppert EE (eds) Microscopic antomy of invertebrates vol 15 Hemichordata Chaetognatha, and the Invertebrate Chordates Wiley-Liss, New York, pp 349-504.

Sahlin K. 1988. Gastrin/CCK-like immunoreactivity in Hatschek's groove of *Branchiostoma lanceolatum* (Cephalochordata). Gen Comp Endocrinol 70(3):436-441.

Sahlin K, Olsson R. 1986. The wheel organand Hatschek's groove in the lancelet, *Branchiostoma lanceolatum* (Cephalochordata). Acta Zool 67(4):201-209.

Salzet M, Bulet P, Van Dorsselaer A, Malecha J. 1993. Isolation, structural characterization and biological function of a lysine-conopressin in the central nervous system of the pharyngobdellid leech *Erpobdella octoculata*. Eur J Biochem 217(3):897-903.

Salzet M, Salzet-Raveillon B, Cocquerelle C, Verger-Bocquet M, Pryor SC, Rialas CM, Laurent V, Stefano GB. 1997. Leech immunocytes contain proopiomelanocortin: nitric oxide mediates hemolymph proopiomelanocortin processing. J Immunol

159(11):5400-5411.

Satake H, Takuwa K, Minakata H, Matsushima O. 1999. Evidence for conservation of the vasopressin/oxytocin superfamily in Annelida. J Biol Chem 274(9):5605-5611.

Sauka-Spengler T, Meulemans D, Jones M, Bronner-Fraser M. 2007. Ancient evolutionary origin of the neural crest gene regulatory network. Developmental cell 13(3):405-420.

Schomerus C, Korf HW, Laedtke E, Moret F, Zhang Q, Wicht H. 2008. Nocturnal behavior and rhythmic period gene expression in a lancelet, *Branchiostoma lanceolatum*. J Biol Rhythms 23(2):170-181.

Schubert M, Yu JK, Holland ND, Escriva H, Laudet V, Holland LZ. 2005. Retinoic acid signaling acts via Hox1 to establish the posterior limit of the pharynx in the chordate amphioxus. Development 132(1):61-73.

Schwind JL. 1928. The development of the hypophysis cerebri of the albino rat. Am J Anat 41:295-315.

Scully KM, Rosenfeld MG. 2002. Pituitary development: regulatory codes in mammalian organogenesis. Science (5563):2231-2235.

Shughrue PJ, Lane MV, Merchenthaler I. 1997. Comparative distribution of estrogen receptor-α and -β mRNA in the rat central nervous system. J Comp Neurol 388(4):507-525.

Simerly RB, Swanson LW, Chang C, Muramatsu M. 1990. Distribution of androgen and estrogen receptor mRNA-containing cells in the rat brain: an *in situ* hybridization study. J Comp Neurol 294(1):76-95.

Smith AI, Funder JW. 1988. Proopiomelanocortin processing in the pituitary, central nervous system, and peripheral tissues. Endocrine Rev 9(1):159.

Sower SA, Moriyama S, Kasahara M, Takahashi A, Nozaki M, Uchida K, Dahlstrom JM, Kawauchi H. 2006. Identification of sea lamprey GTHβ-like cDNA and its evolutionary implications. Gen Comp Endocrinol 148(1):22-32.

Sower SA, Freamat M, Kavanaugh SI. 2009. The origins of the vertebrate hypothalamic-pituitary-gonadal (HPG) and hypothalamic-pituitary-thyroid (HPT) endocrine systems: new insights from lampreys. Gen Comp Endocrinol 161(1):20-29.

Stafflinger E, Hansen KK, Hauser F, Schneider M, Cazzamali G, Williamson M, Grimmelikhuijzen CJ. 2008. Cloning and identification of an oxytocin/vasopressin-like receptor and its ligand from insects. Proc Natl Acad Sci 105(9):3262-3267.

Stefano GB, Salzet-Raveillon B, Salzet M. 1999. *Mytilus edulis* hemolymph contains pro-opiomelanocortin: LPS and morphine stimulate differential processing. Brain Res 63(2):340-350.

Sudo S, Kuwabara Y, Park JI, Hsu SY, Hsueh AJ. 2005. Heterodimeric fly glycoprotein hormone-α2 (GPA2) and glycoprotein hormone-β5 (GPB5) activate fly leucine-rich repeat-containing G protein-coupled receptor-1 (DLGR1) and stimulation of human thyrotropin receptors by chimeric fly GPA2 and human GPB5. Endocrinology 146(8):3596-3604.

Swaab DF, Hofman MA, Lneassen PJ, Purba JS, Raadsheer FC, Van de Nes JAP. 1993. Functional neuroanatomy and neuropathology of the human hypothala. Anatom Embryol 187(4):317-330.

Takuwa-Kuroda K, Iwakoshi-Ukena E, Kanda A, Minakata H. 2003. Octopus, which owns the most advanced brain in invertebrates, has two members of vasopressin/oxytocin superfamily as in vertebrates. Regul Pept 115(2):139-149.

Tateno H, Saneyoshi A, Ogawa T, Muramoto K, Kamiya H, Saneyoshi M. 1998. Isolation and characterization of rhamnose-binding lectins from eggs of steelhead trout (*Oncorhynchus mykiss*) homologous to low density lipoprotein receptor superfamily. J Biol Chem 273(30):19190-19197.

Tello JA, Sherwood NM. 2009. Amphioxus: beginning of vertebrate and end of invertebrate type GnRH receptor lineage. Endocrinology 150(6):2847-2856.

Tessmar-Raible K, Raible F, Christodoulou F, Guy K, Rembold M, Hausen H, Arendt D. 2007. Conserved sensory-neurosecretory cell types in annelid and fish forebrain: insights into hypothalamus evolution. Cell 129(7):1389-1400.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22(22):4673-4680.

Tjoa LT, Welsch U. 1974. Electron microscopical observations on Kolliker's and Hatschek's pit and on the wheel organ in the head region of Amphioxus (*Branchiostoma lanceolatum*). Cell Tissue Res 153(2):175-187.

Tsukahara T, Gorbman A, Kobayashi H. 1986. Median eminence equivalence of the neurohypophysis of the hagfish, *Eptatretus burgeri*. Gen Comp Endocrinol 61(3):348-354.

Uemura H, Tezuka Y, Hasegawa C, Kobayashi H. 1994. Immunohistochemical investigation of neuropeptides in the central nervous system of the amphioxus, *Branchiostoma belcheri*. Cell Tissue Res 277:279-287.

Vitt UA, Hsu SY, Hsueh AJ. 2001. Evolution and classification of cystine knot-containing hormones and related extracellular signaling molecules. Mol Endocrinol 15(5):681-694.

Wada H, Okuyama M, Satoh N, Zhang S. 2006. Molecular evolution of fibrillar collagen in chordates, with implications for the evolution of vertebrate skeletons and chordate phylogeny. Evol Dev 8(4):370-377.

Wada H, Satoh N. 1994. Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. Proc Natl Acad Sci 91(5):1801-1804.

Watakabe A, Ichinohe N, Ohsawa S, Hashikawa T, Komatsu Y, Rockland KS, Yamamori T. 2007. Comparative analysis of layer-specific genes in Mammalian neocortex. Cereb Cortex 17(8):1918-1933.

Watanabe Y, Shiina N, Shinozaki F, Yokoyama H, Kominami J, Nakamura-Tsuruta S, Hirabayashi J, Sugahara K, Kamiya H, Matsubara H, Ogawa T, Muramoto K. 2008.

Isolation and characterization of l-rhamnose-binding lectin, which binds to microsporidian Glugea plecoglossi, from ayu (*Plecoglossus altivelis*) eggs. Dev Comp Immunol 32(5):487-499.

Wells MJ, Wells J. 1969. Pituitary analogue in the octopus. Nature 222(5190):293-294.

Welsch LT, Welsch U. 1978. Histolofische und elektronenmikroskopische Untersuchungen an der präoralen Wimpergrude von *Saccoglossus horsti* (Hemichordata) und der Hatschekschen Grube von *Branchiostoma lanceolatum* (Acrania). Ein Beitrag zur phylogenetischen Entwicklung der Adenohypophyse. Zool Jb Anat 100:564-578.

Wicht H, Lacalli TC. 2005. The nervous system of amphioxus: structure, development, and evolutionary significance. Can J Zool 83:122-150.

Willey A. 1894. Amphioxus and the ancestry of the vertebrates. MacMillan, New York.

Xing Y, Myers RV, Cao D, Lin W, Jiang M, Bernard MP, Moyle WR. 2004. Glycoprotein hormone assembly in the endoplasmic reticulum: IV. Probable mechanism of subunit docking and completion of assembly. J Biol Chem 279(34):35458-35468.

Yasui K, Zhang S, Uemura M, Saiga H. 2000. Left-right asymmetric expression of BbPtx, a Ptx-related gene, in a lancelet species and the developmental left-sidedness in deuterostomes. Development 127(1):187-195.

Yu JK, Satou Y, Holland ND, Shin IT, Kohara Y, Satoh N, Bronner-Fraser M, Holland LZ. 2007. Axial patterning in cephalochordates and the evolution of the organizer. Nature 445(7128):613-617.

Acknowledgement

I express my sincere gratitude to my research supervisor, Prof. Kaoru Kubokawa, Ocean Research Institute, The University of Tokyo, for advice and discussion throughout five years of my study. A lot of chances to do the experiments of my interests and also wonderful experiences on the research cruise which provided by her are invaluable for me.

I am also grateful to Dr. Toyoji Kaneko for beneficial advice to completing my thesis. I am also grateful to Dr. Min Kyun Park, Department of Biological Science, for precise suggestion for my research. I am also grateful to Koji Inoue, for profitable comments for my present study. I am extremely grateful to Dr. Hironori Ando, Kyusyu University, for discussion and excellent technique for construction of Hatschek's pit cDNA library. Without his help, I would not complete EST analysis of Hatschek's pit.

I feel gratitude to Dr. Akihisa Urano for advice from experimental technique to scientific knowledge. I am grateful to Dr. Yoshihito Niimura, Tokyo Medical and Dental University, for genomic survey of the pituitary hormone homologous genes in amphioxus. I am grateful to Dr. Takehiro Kusakabe, Dr. Jae Man Lee, Kyusyu University, for construction recombinant proteins. I express my grateful to Dr. Shinji Nagata, Department of Applied Biological Chemistry, for providing instruments for peptide analysis and helpful technical advice.

I want to express my gratitude to the people of Akabane fishing port for collecting amphioxus. I give my thanks to the amphioxus used for my study.

A special note of thanks to Dr. Takanobu Mizuta, Dr. Yasuyo Shigetani, Dr. Roy Sonali, Ms. Mayumi Inaba, Mr. Hisayuki Iwata, the former members of the laboratory, for their advice, encouragement, and friendship. Many thanks to the members of Ocean Research Institute and associates I have met during my research. Special thanks to Ms. Tomoko Koito for friendship.

Finally, I would like to record my appreciation of my parents and my sister for supporting me all the time.

Appendix. Result of Blast search of the clones obtained by EST analysis of Hatschek's pit cDNA library. The clones which matched with NCBI non-redundant data are listed in order of E-value.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
1	Contig90	4	ref|XP_862471.1| PREDICTED: similar to tubulin, alpha 1 isoform 10 [Canis familiaris]	0.00E+00	102751| alpha-tubulin
2	Contig114	8	ref|NP_776276.1| cytochrome c oxidase subunit I [B. belcheri]	0.00E+00	No hit
3	Contig20	3	dbj|BAC75528.1| NADH dehydrogenase subunit 5 [B. belcheri]	0.00E+00	No hit
4	Contig25	4	dbj|BAH86105.1| cytochrome b [B. belcheri]	4.00E-159	No hit
5	Contig19	2	ref|XP_002613759.1| hypothetical protein BRAFLDRAFT_84515 [B. floridae]	2.00E-144	84515| guanine nucreotide-binding protein
6	Contig126	7	ref|XP_002593183.1| hypothetical protein BRAFLDRAFT_277869 [B. floridae]	6.00E-142	277869| similar to ribosomal protein L8
7	Contig133	2	gb|AAQ18147.1| 14-3-3 protein [B. belcheri]	1.00E-127	254312| 14-3-3 protein
8	Contig34	9	ref|NP_776277.1| cytochrome c oxidase subunit II [B. belcheri]	1.00E-117	No hit
9	Contig78	4	ref|XP_002587985.1| hypothetical protein BRAFLDRAFT_125387 [B. floridae]	3.00E-115	125387| ribosomal protein L7
10	Contig53	4	gb|AAN52373.1| ribosomal protein L15 [B. belcheri]	1.00E-104	271350| ribosomal protein L15
11	32-68_TriplEx25	1	ref|XP_002588504.1| hypothetical protein BRAFLDRAFT_220697 [B. floridae]	4.00E-100	220697| guanine nucleotide-binding protein beta subunit 1
12	1-65_TriplEx25	1	gb|EFA07008.1| hypothetical protein TcasGA2_TC009978 [Tribolium castaneum]	7.00E-100	113894| similar to ubiquitin C isoform 21
13	Contig17	4	ref|XP_002608162.1| hypothetical protein BRAFLDRAFT_125871 [B. floridae]	5.00E-99	125871| dermatopontin
14	Contig26	4	gb|AAO31769.1| ribosomal protein L10 [B. belcheri]	9.00E-96	275426| ribosomal protein L10
15	7-74_TriplEx25	1	ref|XP_002595746.1| hypothetical protein BRAFLDRAFT_200833 [B. floridae]	1.00E-94	227398| 40S ribosomal protein SA
16	30-63_TriplEx25	1	gb|ABB86551.1| proteosome PSMB6/9 protein [B. lanceolatum]	7.00E-92	114609| proteasome
17	17-73_TriplEx25	1	gb|AAN52383.1| ribosomal protein L9 [B. belcheri]	1.00E-91	267499|ribosomal protein L9
18	Contig166	2	ref|XP_002611830.1| hypothetical protein BRAFLDRAFT_114734 [B. floridae]	2.00E-89	116135| MGC84358 protein
19	Contig153	2	hypothetical protein BRAFLDRAFT_126054 [B. floridae]	2.00E-89	126054| 80906 protein
20	20-25_TriplEx25	1	ref|XP_002607766.1| hypothetical protein BRAFLDRAFT_123247 [B. floridae]	4.00E-89	199520| similar to Phosphatidylethanolamine-Binding protein
21	Contig132	7	gb|AAQ96652.1| ribosomal protein L17 [B. belcheri]	1.00E-88	124327| ribosomal protein L17
22	Contig168	5	ref|XP_002599076.1| hypothetical protein BRAFLDRAFT_81739 [B. floridae]	1.00E-87	81739| ribosomal protein S4
23	40-56_TriplEx25	1	ref|XP_002608922.1| hypothetical protein BRAFLDRAFT_85513 [B. floridae]	8.00E-87	257602| nucleolar protein 5A
24	Contig142	3	dbj|BAH86164.1| cytochrome c oxidase subunit III [B. belcheri]	9.00E-87	No hit
25	Contig50	2	gb|AAN52374.1| ribosomal protein L18a [B. belcheri]	1.00E-86	287192| ribosomal protein L18a
26	36-20_TriplEx25	1	ref|XP_002595117.1| hypothetical protein BRAFLDRAFT_113766 [B. floridae]	2.00E-86	113766| ATP synthase beta chain, mitochondrial precursor
27	17-35_TriplEx25	1	ref|XP_002596367.1| hypothetical protein BRAFLDRAFT_76176 [B. floridae]	5.00E-86	109403| protein kinase
28	3-71_TriplEx25	1	ref|XP_002603622.1| hypothetical protein BRAFLDRAFT_115449 [B. floridae]	1.00E-84	115449| KDEL endoplasmic reticulum protein retention reseptor 2
29	16-67_TriplEx25	1	gb|AAQ96656.1| adenosylhomocysteinase [B. belcheri]	4.00E-83	113595| adenosylhomocysteinase
30	4-75_TriplEx25	1	ref|XP_002594490.1| hypothetical protein BRAFLDRAFT_124962 [B. floridae]	8.00E-83	124962| similar to melanotransferrin precursor
31	30-40_TriplEx25	1	ref|XP_002610692.1| hypothetical protein BRAFLDRAFT_117936 [B. floridae]	4.00E-82	117936| translation initiation factor
32	20-37_TriplEx25	1	ref|XP_002608162.1| hypothetical protein BRAFLDRAFT_125871 [B. floridae]	8.00E-82	125871| dermatopontin
33	26-26_TriplEx25	1	emb|CAO98872.1| hypothetical protein [Nakaseomyces delphensis]	4.00E-81	285403| ribosomal protein L27a
34	Contig103	2	gb|AAP93925.1| phosphatidylethanolamine-binding protein [B. belcheri]	1.00E-79	103077| phosphatidylethanolamine-binding protein
35	Contig129	2	dbj|BAH86146.1| NADH dehydrogenase subunit 2 [B. belcheri]	2.00E-79	No hit
36	16-46_TriplEx25	1	gb|AAM18861.1|AF391287_2 unknown [B. floridae]	4.00E-79	186174| spliceosome RNA helicase BAT1
37	38-59_TriplEx25	1	ref|XP_002587561.1| hypothetical protein BRAFLDRAFT_282752 [B. floridae]	6.00E-79	282752| Na+/solute symporter
38	3-22_TriplEx25	1	gb|AAO31771.1| ribosomal protein L21 [B. belcheri]	2.00E-78	125590| ribosomal protein L21
39	26-29_TriplEx25	1	ref|XP_002594706.1| hypothetical protein BRAFLDRAFT_285448 [B. floridae]	3.00E-78	281254| hypothytical protein
40	3-25_TriplEx25	1	ref|XP_002591842.1| hypothetical protein BRAFLDRAFT_125332 [B. floridae]	5.00E-78	287506| ribosomal protein L12
41	27-44_TriplEx25	1	dbj|BAH86193.1| NADH dehydrogenase subunit 4 [B. belcheri]	3.00E-77	No hit
42	Contig141	3	gb|AAM09534.1|AF491451_1 ribosomal protein S19 [B. belcheri]	4.00E-77	126768| 40S ribosomal protein S19
43	34-16_TriplEx25	1	ref|XP_002608242.1| hypothetical protein BRAFLDRAFT_125057 [B. floridae]	7.00E-77	125057| mucin-5B precursor
44	Contig9	2	ref|XP_002603023.1| hypothetical protein BRAFLDRAFT_123990 [B. floridae]	3.00E-76	123990| ATP synthase c-subunit precursor
45	25-74_TriplEx25	1	gb|AAN52388.1| ribosomal protein S16 [B. belcheri]	3.00E-76	122665| ribosomal protein S16
46	22-88_TriplEx25	1	gb|AAN52387.1| ribosomal protein S13 [B. belcheri]	5.00E-76	115249| ribosomal protein S13
47	Contig1	7	gb|AAN86978.1| ribosomal protein S23 [B. belcheri]	3.00E-75	287235| ribosomal protein S23
48	4-63_TriplEx25	1	ref|XP_001166059.1| PREDICTED: similar to OCP-II protein [Pan troglodytes]	1.00E-74	124598| S-phase kinase-associated
49	28-38_TriplEx25	1	ref|XP_002605825.1| hypothetical protein BRAFLDRAFT_84310 [B. floridae]	1.00E-73	9085| ATP-gated cation channel receptor P2X4
50	27-48_TriplEx25	1	ref|XP_002608309.1| hypothetical protein BRAFLDRAFT_89286 [B. floridae]	7.00E-73	89286| hypothetical protein
51	48-20_TriplEx25	1	ref|XP_002595381.1| hypothetical protein BRAFLDRAFT_119005 [B. floridae]	1.00E-72	117458| hypothetical protein
52	Contig83	9	ref|XP_002610900.1| hypothetical protein BRAFLDRAFT_91496 [B. floridae]	2.00E-72	91496| similar to proprotein convertase 6B
53	49-33_TriplEx25	1	ref|XP_002612994.1| hypothetical protein BRAFLDRAFT_120829 [B. floridae]	5.00E-72	120829| thymocyte nuclear protein 1
54	19-89_TriplEx25	1	ref|XP_002605229.1| hypothetical protein BRAFLDRAFT_126601 [B. floridae]	5.00E-72	126601| stromal cell derived factor 2-like protein
55	25-42_TriplEx25	1	ref|XP_002593962.1| hypothetical protein BRAFLDRAFT_68601 [B. floridae]	1.00E-71	68601| VAMP (vesicle-associated membrane protein)-associated protein
56	9-38_TriplEx25	1	ref|XP_002608931.1| hypothetical protein BRAFLDRAFT_124233 [B. floridae]	2.00E-71	124233| isocitrate dehydrogenase subunit beta
57	25-89_TriplEx25	1	ref|XP_002604169.1| hypothetical protein BRAFLDRAFT_120398 [B. floridae]	3.00E-71	120398| similar to EGF-like repeats discoidin I-like domains
58	Contig95	5	ref|XP_002591268.1| hypothetical protein BRAFLDRAFT_121421 [B. floridae]	5.00E-71	121421| hypothetical protein
59	15-19_TriplEx25	1	ref|XP_002602109.1| hypothetical protein BRAFLDRAFT_98958 [B. floridae]	1.00E-70	98958| mannan endo-1,4-beta-mannosidase precursor
60	17-54_TriplEx25	1	ref|XP_002588121.1| hypothetical protein BRAFLDRAFT_124949 [B. floridae]	3.00E-70	124949| Pr2 protein
61	9-64_TriplEx25	1	ref|XP_002596647.1| hypothetical protein BRAFLDRAFT_280235 [B. floridae]	5.00E-70	289371| protein phosphatase
62	9-62_TriplEx25	1	ref|XP_002595531.1| hypothetical protein BRAFLDRAFT_130427 [B. floridae]	2.00E-69	200215| actin related protein 2/3 complex
63	Contig150	2	gb|AAS91553.1| AmphiHMG1/2 [B. belcheri]	2.00E-69	57556| high mobility group
64	28-54_TriplEx25	1	ref|XP_002608790.1| hypothetical protein BRAFLDRAFT_125595 [B. floridae]	3.00E-69	125595| alphaP integrin
65	Contig28	4	ref|XP_002609941.1| hypothetical protein BRAFLDRAFT_114927 [B. floridae]	9.00E-69	124370| ribosomal protein S18
66	40-84_TriplEx25	1	ref|XP_002607973.1| hypothetical protein BRAFLDRAFT_74922 [B. floridae]	2.00E-68	120952| similar to cytochrome P450, family 3, subfamily a, polypeptide 1
67	48-33_TriplEx25	1	ref|XP_002599282.1| hypothetical protein BRAFLDRAFT_64360 [B. floridae]	2.00E-68	64360 etc/casein kinase
68	17-80_TriplEx25	1	gb|AAQ96656.1| adenosylhomocysteinase [B. belcheri]	1.00E-67	113595| adenosylhomocysteinase
69	Contig2	2	gb|AAK91296.1|AF395864_1 ubiquitin [B. belcheri]	2.00E-67	113894| similar to ubiquitin C isoform 21
70	17-49_TriplEx25	1	ref|XP_002604725.1| hypothetical protein BRAFLDRAFT_222394 [B. floridae]	4.00E-67	222394| unnamed protein product
71	20-75_TriplEx25	1	ref|XP_002592789.1| hypothetical protein BRAFLDRAFT_275667 [B. floridae]	1.00E-66	275667| Intraflagellar transport 80 homolog
72	28-42_TriplEx25	1	ref|XP_002609424.1| hypothetical protein BRAFLDRAFT_114978 [B. floridae]	3.00E-66	114978| similar to 26S proteasome non-ATPase regulatory subunit 5
73	36-54_TriplEx25	1	ref|XP_002587677.1| hypothetical protein BRAFLDRAFT_116897 [B. floridae]	1.00E-65	116897| type I keratin k1
74	Contig39	4	gb|AAN86979.1| ribosomal protein S15a [B. belcheri]	2.00E-65	128055| ribosomal protein S15a
75	27-52_TriplEx25	1	ref|XP_002604974.1| hypothetical protein BRAFLDRAFT_126703 [B. floridae]	3.00E-65	126703| hyothetical protein
76	38-46_TriplEx25	1	gb|AAQ83886.1| DC2-like protein [B. belcheri]	4.00E-65	273440| hypothetical protein
77	41-36_TriplEx25	1	ref|XP_002592201.1| hypothetical protein BRAFLDRAFT_123946 [B. floridae]	5.00E-65	125128| ribosomal protein L23
78	Contig67	2	gb|AAT45380.1| apextrin [B. belcheri]	5.00E-65	126750| hypothetical protein
79	42-55_TriplEx25	1	ref|XP_002603275.1| hypothetical protein BRAFLDRAFT_281712 [B. floridae]	8.00E-64	281712| similar to Prefoldin subunit 5
80	25-78_TriplEx25	1	ref|XP_002609380.1| hypothetical protein BRAFLDRAFT_124610 [B. floridae]	1.00E-63	128859| similar to collagen, type XXII, alpha 1
81	37-90_TriplEx25	1	ref|NP_776277.1| cytochrome c oxidase subunit II [B. belcheri]	4.00E-63	No hit
82	15-6_TriplEx25	1	ref|XP_002606009.1| hypothetical protein BRAFLDRAFT_129508 [B. floridae]	8.00E-63	126543| NADH-ubiquinone oxidoreductase
83	9-56_TriplEx25	1	ref|XP_002611705.1| hypothetical protein BRAFLDRAFT_117078 [B. floridae]	4.00E-62	289131| 60S ribosomal protein L31
84	40-37_TriplEx25	1	ref|XP_002605916.1| hypothetical protein BRAFLDRAFT_87410 [B. floridae]	1.00E-61	110869| similar to fatty acid synthase
85	Contig128	2	ref|XP_002594865.1| hypothetical protein BRAFLDRAFT_124448 [B. floridae]	1.00E-61	124448| similar to proteasome subunit beta type 1
86	40-96_TriplEx25	1	ref|XP_002603018.1| hypothetical protein BRAFLDRAFT_114843 [B. floridae]	5.00E-61	130017| heat shock cognate 71 kDa protein
87	38-21_TriplEx25	1	gb|AAI65718.1| Rars protein [Danio rerio]	5.00E-61	289113| arginyl-tRNA synthetase
88	4-71_TriplEx25	1	ref|XP_002592147.1| hypothetical protein BRAFLDRAFT_114866 [B. floridae]	7.00E-61	271557| /ribosomal protein L32
89	32-84_TriplEx25	1	ref|XP_002610036.1| hypothetical protein BRAFLDRAFT_237984 [B. floridae]	1.00E-60	237951| similar to TNF receptor associated factor 3
90	22-28_TriplEx25	1	ref|XP_002608642.1| hypothetical protein BRAFLDRAFT_172198 [B. floridae]	2.00E-60	172198| cytochrome c oxidase, subunit Vb
91	Contig62	2	gb|AAM28852.1|AF503586_1 ribosomal protein S20 [B. belcheri]	6.00E-60	284108| ribosomal protein S20
92	3-64_TriplEx25	1	ref|XP_002588598.1| hypothetical protein BRAFLDRAFT_107521 [B. floridae]	1.00E-59	107521| casein kinase I isoform alpha
93	Contig63	3	ref|NP_776274.1| NADH dehydrogenase subunit 1 [B. belcheri]	1.00E-59	No hit
94	Contig117	2	ref|XP_002611705.1| hypothetical protein BRAFLDRAFT_117078 [B. floridae]	4.00E-59	289131| 60S ribosomal protein L31
95	28-65_TriplEx25	1	gb|AAO31772.1| ribosomal protein L34 [B. belcheri]	4.00E-59	276111| ribosomal protein L34
96	28-39_TriplEx25	1	gb|AAM18076.1|AF498232_1 peroxiredoxin V protein [B. belcheri]	4.00E-59	119799| peroxiredoxin V protein
97	Contig155	7	ref|XP_002610900.1| hypothetical protein BRAFLDRAFT_91496 [B. floridae]	8.00E-59	121050| hyoithetical protein
98	15-41_TriplEx25	1	ref|XP_002609766.1| hypothetical protein BRAFLDRAFT_122094 [B. floridae]	2.00E-58	122094| hyothetical protein
99	Contig119	2	ref|XP_002608707.1| hypothetical protein BRAFLDRAFT_58108 [B. floridae]	2.00E-58	60290| ribosomal protein L11
100	32-22_TriplEx25	1	ref|XP_002598175.1| hypothetical protein BRAFLDRAFT_143676 [B. floridae]	2.00E-58	143676| FRAS1-related extracellular

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
101	6-57_TriplEx25	1	gb|ABB85359.1| Ran [B. belcheri]	2.00E-58	282949| ran
102	5-24_TriplEx25	1	ref|XP_002605352.1| hypothetical protein BRAFLDRAFT_212636 [B. floridae]	8.00E-58	212636| prostaglandin-D synthase
103	48-4_TriplEx25	1	sp|Q86QN6.1|BDEF_BRABE RecName: Full=Big defensin [B. belcheri]	1.00E-57	No hit
104	22-45_TriplEx25	1	ref|XP_002611790.1| hypothetical protein BRAFLDRAFT_268894 [B. floridae]	1.00E-57	288998| DnaJ homolog subfamily C member 8
105	Contig113	3	gb|AAT39881.1| ribosomal protein S15 [B. belcheri]	2.00E-57	99672| ribosomal protein S19/S15
106	Contig3	5	gb|ABK27629.1| ribosomal protein S30 [B. belcheri]	7.00E-57	116267| ribosomal protein S30
107	49-66_TriplEx25	1	ref|XP_002595826.1| hypothetical protein BRAFLDRAFT_115715 [B. floridae]	8.00E-57	115715| transaldolase 1
108	27-42_TriplEx25	1	ref|XP_002602335.1| hypothetical protein BRAFLDRAFT_268377 [B. floridae]	2.00E-56	98035| similar to 40S ribosomal protein S8
109	22-79_TriplEx25	1	ref|XP_002612233.1| hypothetical protein BRAFLDRAFT_284837 [B. floridae]	3.00E-56	284837| adenosine/AMP deaminase
110	36-32_TriplEx25	1	ref|XP_002591791.1| hypothetical protein BRAFLDRAFT_123537 [B. floridae]	3.00E-56	123537| tektin-2
111	37-14_TriplEx25	1	gb|AAO18673.1| signal recognition particle 19 kDa protein [B. belcheri]	3.00E-56	268468| signal recognition particle 19 kDa protein
112	24-17_TriplEx25	1	ref|XP_002593769.1| hypothetical protein BRAFLDRAFT_245767 [B. floridae]	4.00E-56	245767| hypothetical protein
113	8-10_TriplEx25	1	ref|XP_002595969.1| hypothetical protein BRAFLDRAFT_60971 [B. floridae]	5.00E-56	60971| arp2/3 complex 21kDa subunit
114	33-89_TriplEx25	1	gb|AAO18674.1| defensin [B. belcheri]	5.00E-56	No hit
115	34-9_TriplEx25	1	gb|AAN73381.1| ribosomal protein L18 [B. lanceolatum]	5.00E-56	275280| ribosomal protein L18
116	46-83_TriplEx25	1	sp|Q86QN6.1|BDEF_BRABE RecName: Full=Big defensin [B. belcheri]	3.00E-55	No hit
117	34-86_TriplEx25	1	ref|XP_002202064.1| hypothetical protein BRAFLDRAFT_64586 [B. floridae]	6.00E-55	64586| three prime repair exonuclease 1 isoform b
118	33-1_TriplEx25	1	ref|XP_002613946.1| hypothetical protein BRAFLDRAFT_67491 [B. floridae]	8.00E-54	67491| SMART/HDAC1 associated repressor protein
119	Contig82	2	ref|XP_002595401.1| hypothetical protein BRAFLDRAFT_69231 [B. floridae]	3.00E-53	69231| SEC14-like 1
120	33-63_TriplEx25	1	ref|XP_002586803.1| hypothetical protein BRAFLDRAFT_270722 [B. floridae]	4.00E-53	270722| NADH dehydrogenase
121	48-72_TriplEx25	1	ref|XP_002612554.1| hypothetical protein BRAFLDRAFT_122184 [B. floridae]	7.00E-53	122184| D-amino acid oxidase 1
122	37-24_TriplEx25	1	ref|XP_002592029.1| hypothetical protein BRAFLDRAFT_122398 [B. floridae]	1.00E-52	122398| ribosomal protein L27e
123	19-40_TriplEx25	1	ref|XP_002608802.1| hypothetical protein BRAFLDRAFT_125603 [B. floridae]	5.00E-52	115322l eongation factor 1-beta
124	47-85_TriplEx25	1	gb|AAN52378.1| ribosomal protein L26 [B. belcheri]	5.00E-52	263523l ribosomal protein L26
125	4-90_TriplEx25	1	ref|XP_002588872.1| hypothetical protein BRAFLDRAFT_235909 [B. floridae]	9.00E-52	268762l mitogen-activated protein kinase 1
126	30-90_TriplEx25	1	ref|XP_002613895.1| hypothetical protein BRAFLDRAFT_277494 [B. floridae]	2.00E-51	277494l translation initiation factor 4 gamma
127	36-21_TriplEx25	1	gb|ACI66768.1| 40S ribosomal protein S7 [Salmo salar]	3.00E-51	288119l ribosomal protein S7
128	Contig44	2	ref|XP_002598366.1| hypothetical protein BRAFLDRAFT_69722 [B. floridae]	4.00E-51	69722l 40S ribosomal protein S3a
129	Contig157	3	ref|XP_002604169.1| hypothetical protein BRAFLDRAFT_120398 [B. floridae]	6.00E-51	120398l EGF-like repeats and discoidin I like domains-containing protein
130	42-66_TriplEx25	1	ref|XP_002608242.1| hypothetical protein BRAFLDRAFT_125057 [B. floridae]	7.00E-51	125057l mucin-5B precursor
131	Contig10	9	ref|XP_002608162.1| hypothetical protein BRAFLDRAFT_125871 [B. floridae]	8.00E-51	125871l dermatopontin
132	14-36_TriplEx25	1	gb|AAT39882.1| F1Fo-ATPase synthase f subunit [B. belcheri]	3.00E-50	126066l ATPase syntase f subunit
133	47-87_TriplEx25	1	ref|XP_002200475.1| serologically defined colon cancer antigen [Taeniopygia guttata]	8.00E-50	129091l serologically defined colon cancer antigen 1
134	23-59_TriplEx25	1	ref|XP_002607146.1| hypothetical protein BRAFLDRAFT_118658 [B. floridae]	1.00E-49	123165l hypothetical protein
135	Contig158	2	gb|AAT39882.1| F1Fo-ATPase synthase f subunit [B. belcheri]	1.00E-49	126066l hypothetical protein
136	1-34_TriplEx25	1	gb|AAN52381.1| ribosomal protein L36 [B. belcheri]	5.00E-49	279545l ribosomal protein L36
137	27-85_TriplEx25	1	ref|XP_002599178.1| hypothetical protein BRAFLDRAFT_275202 [B. floridae]	6.00E-49	275202l COP9 signalosome complex subunit 3
138	Contig15	2	ref|XP_002586263.1| hypothetical protein BRAFLDRAFT_273508 [B. floridae]	1.00E-48	288811l 60S ribosomal protein L24
139	Contig122	2	ref|XP_002608161.1| hypothetical protein BRAFLDRAFT_125872 [B. floridae]	2.00E-48	125872l MGC85507 protein
140	25-59_TriplEx25	1	ref|XP_002590150.1| hypothetical protein BRAFLDRAFT_126055 [B. floridae]	6.00E-48	275430l ribosomal protein L3
141	25-5_TriplEx25	1	ref|XP_002588236.1| hypothetical protein BRAFLDRAFT_86681 [B. floridae]	6.00E-48	86681l hyoithetical protein
142	Contig43	2	gb|AAT39418.1| rhamnose-binding lectin precursor [B. belcheri]	1.00E-47	110271l rhamnose-binding lectin precursor
143	Contig24	5	gb|AAN52380.1| ribosomal protein L35 [B. belcheri]	2.00E-47	82089l ribosomal protein L35
144	27-79_TriplEx25	1	ref|XP_002606048.1| hypothetical protein BRAFLDRAFT_269803 [B. floridae]	4.00E-47	269803l 60S ribosomal protein L37-A
145	Contig4	2	ref|XP_002593397.1| hypothetical protein BRAFLDRAFT_277091 [B. floridae]	6.00E-47	277091l ribosomal protein L35a
146	34-28_TriplEx25	1	gb|AAQ83893.2| glutathione S-transferase [B. belcheri]	6.00E-47	214367l glutathione S-transferase
147	Contig104	5	dbj|BAB97379.1| amyloid protein A [B. belcheri]	6.00E-47	124064l amyloid protein A
148	1-85_TriplEx25	1	gb|AAN52380.1| ribosomal protein L35 [B. belcheri]	7.00E-47	82089l ribosomal protein L35
149	13-81_TriplEx25	1	ref|XP_002613332.1| hypothetical protein BRAFLDRAFT_68299 [B. floridae]	2.00E-46	68299l hypothetical protein
150	47-53_TriplEx25	1	ref|XP_002599076.1| hypothetical protein BRAFLDRAFT_81739 [B. floridae]	3.00E-46	81739l ribosomal protein S4
151	37-76_TriplEx25	1	gb|AAP21779.1| ribosomal protein L36a [B. belcheri]	6.00E-46	92142l ribosomal protein L36A
152	42-2_TriplEx25	1	ref|XP_002587523.1| hypothetical protein BRAFLDRAFT_284483 [B. floridae]	9.00E-46	284483l splicing factor 3b, subunit 5, 10kDa
153	15-72_TriplEx25	1	dbj|BAC75525.1| NADH dehydrogenase subunit 3 [B. belcheri]	1.00E-45	No hit
154	13-48_TriplEx25	1	ref|XP_002591513.1| hypothetical protein BRAFLDRAFT_131051 [B. floridae]	2.00E-45	112860l hypothetical protein
155	Contig7	6	gb|AAP21779.1| ribosomal protein L36a [B. belcheri]	2.00E-45	92142l ribosomal protein L36A
156	35-80_TriplEx25	1	ref|XP_001185846.1| PREDICTED: similar to ankyrin [S. purpuratus]	4.00E-45	117088l similar to ankyrin
157	Contig87	3	gb|AAT39418.1| rhamnose-binding lectin precursor [B. belcheri]	1.00E-44	110271l rhamnose-binding lectin precursor
158	48-14_TriplEx25	1	ref|XP_002612609.1| hypothetical protein BRAFLDRAFT_122160 [B. floridae]	2.00E-44	132593l hypothetical protein Glutatione S-transferase
159	28-2_TriplEx25	1	ref|XP_002606860.1| hypothetical protein BRAFLDRAFT_115348 [B. floridae]	2.00E-44	129027l calreticulin, like 2
160	1-12_TriplEx25	1	dbj|BAH86163.1| ATP synthase F0 subunit 6 [B. belcheri]	3.00E-44	No hit
161	47-26_TriplEx25	1	sp|Q04948.1| non-neuronal cytoplasmic intermediate filament [B. lanceolatum]	4.00E-44	77526l non-neuronal cytoplasmic intermediate filament protein
162	30-19_TriplEx25	1	ref|XP_002607175.1| hypothetical protein BRAFLDRAFT_68030 [B. floridae]	9.00E-44	68030l MGC82029 protein
163	36-92_TriplEx25	1	ref|XP_002604469.1| hypothetical protein BRAFLDRAFT_79230 [B. floridae]	1.00E-43	79230l voltage-dependent T-type calcium channel alpha-1H subunit
164	8-96_TriplEx25	1	ref|XP_002613121.1| hypothetical protein BRAFLDRAFT_210607 [B. floridae]	2.00E-43	210607l amyloid beta A4 protein precursor
165	36-24_TriplEx25	1	ref|XP_002608707.1| hypothetical protein BRAFLDRAFT_58108 [B. floridae]	2.00E-43	60290l ribosomal protein L11
166	19-79_TriplEx25	1	ref|XP_002608166.1| hypothetical protein BRAFLDRAFT_90419 [B. floridae]	2.00E-43	90419 l kunitz-type proteinase inhibitor
167	Contig167	2	ref|XP_002596031.1| hypothetical protein BRAFLDRAFT_66238 [B. floridae]	2.00E-43	66238l ribosomal protein S24
168	36-22_TriplEx25	1	ref|XP_002591268.1| hypothetical protein BRAFLDRAFT_121421 [B. floridae]	3.00E-43	12142l hyoithetical protein
169	Contig108	2	ref|XP_002599767.1| hypothetical protein BRAFLDRAFT_261371 [B. floridae]	4.00E-43	261371l hyoithetical protein
170	29-13_TriplEx25	1	ref|XP_001629285.1| predicted protein [Nematostella vectensis]	3.00E-42	No hit
171	Contig72	3	gb|AAO18674.1| defensin [B. belcheri]	3.00E-42	No hit
172	40-79_TriplEx25	1	ref|XP_002606299.1| hypothetical protein BRAFLDRAFT_67539 [B. floridae]	4.00E-42	133687l MGC53078 protein
173	Contig115	2	ref|XP_002594057.1| hypothetical protein BRAFLDRAFT_68503 [B. floridae]	5.00E-42	68503l bacteridal permability-increasing protein
174	Contig61	3	gb|AAP21828.1| ribosomal protein S21 [B. belcheri]	7.00E-42	124356l 40S ribosomal protein S21
175	15-20_TriplEx25	1	ref|XP_002610006.1| hypothetical protein BRAFLDRAFT_131114 [B. floridae]	1.00E-41	131114l nuclear cap binding protein subunit 1, 80kDa
176	17-59_TriplEx25	1	ref|XP_002589182.1| hypothetical protein BRAFLDRAFT_74660 [B. floridae]	1.00E-41	74660l UPF0338 protein FLJ27310 homolog
177	17-41_TriplEx25	1	ref|XP_002611407.1| hypothetical protein BRAFLDRAFT_117231 [B. floridae]	2.00E-41	117231l sterol regulatory element
178	38-89_TriplEx25	1	ref|XP_002594727.1| hypothetical protein BRAFLDRAFT_114633 [B. floridae]	3.00E-41	272855l translin
179	1-38_TriplEx25	1	ref|XP_002607328.1| hypothetical protein BRAFLDRAFT_261238 [B. floridae]	4.00E-41	261238l nuclear protein
180	18-81_TriplEx25	1	ref|XP_002589599.1| hypothetical protein BRAFLDRAFT_122926 [B. floridae]	1.00E-40	128954l hypothetical protein
181	17-55_TriplEx25	1	ref|XP_002589300.1| hypothetical protein BRAFLDRAFT_268096 [B. floridae]	2.00E-40	268096l small nucear ribonucleoprotein
182	Contig169	4	gb|AAL79538.1|AF470687_1 40S ribosomal protein S12 [B. belcheri]	3.00E-40	109012l 40S ribosomal protein S12
183	Contig110	3	gb|AAK52799.2|AF363029_1 60S ribosomal protein L37A [B. belcheri]	4.00E-40	284457l ribosomal L37ae protein
184	27-39_TriplEx25	1	ref|XP_002603969.1| hypothetical protein BRAFLDRAFT_71743 [B. floridae]	5.00E-40	102346l similar to zinc finger protein 91
185	25-49_TriplEx25	1	ref|XP_002600004.1| hypothetical protein BRAFLDRAFT_74122 [B. floridae]	8.00E-40	129891l lipoxygenase
186	20-20_TriplEx25	1	ref|XP_002592820.1| hypothetical protein BRAFLDRAFT_194454 [B. floridae]	1.00E-39	108228l similar to Zinc finger protein 208 isoform 21
187	1-76_TriplEx25	1	gb|AAN52375.1| ribosomal protein L22 [B. belcheri]	1.00E-39	171706l 60S ribosomal protein L22
188	30-29_TriplEx25	1	ref|XP_002607322.1| hypothetical protein BRAFLDRAFT_119205 [B. floridae]	2.00E-39	124826l 26S proteasome subunit P45
189	50-60_TriplEx25	1	ref|XP_002606996.1| hypothetical protein BRAFLDRAFT_200793 [B. floridae]	2.00E-39	226898l ubiquitin
190	40-5_TriplEx25	1	gb|ACO14423.1| Transmembrane emp24 domain-containing protein [Esox lucius]	2.00E-39	112815l similar to ribokinase; 5230400M11Rik
191	38-8_TriplEx25	1	ref|XP_002612598.1| hypothetical protein BRAFLDRAFT_143394 [B. floridae]	3.00E-39	143394l SH3 domain binding glutamic acid-rich protein like 2
192	17-47_TriplEx25	1	ref|XP_002607158.1| hypothetical protein BRAFLDRAFT_118651 [B. floridae]	3.00E-39	118651l similar to forkhead box J1
193	38-80_TriplEx25	1	ref|XP_002169001.1| PREDICTED: similar to predicted protein [Hydra magnipapillata]	3.00E-39	87107l hypothetical protein
194	17-95_TriplEx25	1	ref|XP_002593398.1| hypothetical protein BRAFLDRAFT_119559 [B. floridae]	4.00E-39	119559l Protein PTDSR
195	24-61_TriplEx25	1	ref|XP_002591883.1| hypothetical protein BRAFLDRAFT_125525 [B. floridae]	7.00E-39	125525l hypothetical protein
196	2-85_TriplEx25	1	ref|XP_002605805.1| hypothetical protein BRAFLDRAFT_123848 [B. floridae]	1.00E-38	123848l similar to [Segment 1 of 2]
197	8-62_TriplEx25	1	ref|XP_002604546.1| hypothetical protein BRAFLDRAFT_122339 [B. floridae]	1.00E-38	122339l EF-hand containing
198	48-24_TriplEx25	1	ref|XP_002595511.1| hypothetical protein BRAFLDRAFT_118960 [B. floridae]	1.00E-38	118960l hypothetical protein
199	25-76_TriplEx25	1	ref|XP_002597557.1| hypothetical protein BRAFLDRAFT_82338 [B. floridae]	2.00E-38	82338l tumor necrosis factor receptor superfamily, member 11a
200	29-8_TriplEx25	1	ref|XP_002592393.1| hypothetical protein BRAFLDRAFT_67258 [B. floridae]	2.00E-38	67258l hypothetical protein XP_782717

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
201	13-19_TriplEx25	1	ref|XP_002613212.1| hypothetical protein BRAFLDRAFT_278055 [B. floridae]	3.00E-38	278055| heat shock protein 90
202	46-34_TriplEx25	1	ref|XP_002605925.1| hypothetical protein BRAFLDRAFT_87401 [B. floridae]	1.00E-37	110875| Pulmonary surfactant-associated protein C precursor
203	35-15_TriplEx25	1	ref|XP_002601286.1| hypothetical protein BRAFLDRAFT_81325 [B. floridae]	1.00E-37	81325| ATP synthase g
204	23-85_TriplEx25	1	ref|XP_002594638.1| hypothetical protein BRAFLDRAFT_264495 [B. floridae]	1.00E-37	264495| matrix metalloproteinase 14
205	Contig33	14	ref|XP_002596037.1| hypothetical protein BRAFLDRAFT_118076 [B. floridae]	2.00E-37	132915| zeta&apos globin
206	19-18_TriplEx25	1	ref|XP_002595387.1| hypothetical protein BRAFLDRAFT_69217 [B. floridae]	4.00E-37	69217| DNA ligase I
207	17-5_TriplEx25	1	ref|XP_002612507.1| hypothetical protein BRAFLDRAFT_75371 [B. floridae]	7.00E-37	116238| glutathione S-transferase
208	41-46_TriplEx25	1	ref|XP_002613292.1| hypothetical protein BRAFLDRAFT_68257 [B. floridae]	5.00E-36	68257| hypothetical protein
209	40-86_TriplEx25	1	gb|AAK84394.1|AF397146_1 translationally-controlled tumor protein [B. belcheri]	7.00E-36	117727| translationally-controlled tumor protein
210	18-38_TriplEx25	1	ref|XP_002608242.1| hypothetical protein BRAFLDRAFT_125057 [B. floridae]	8.00E-36	125057| mucin-5B precursor
211	1-20_TriplEx25	1	ref|XP_002587089.1| hypothetical protein BRAFLDRAFT_285972 [B. floridae]	2.00E-35	285972| Bloc1s2-prov protein
212	18-55_TriplEx25	1	gb|AAO31776.1| ribosomal protein S10 [B. belcheri]	2.00E-35	125059| ribosomal protein S10
213	6-67_TriplEx25	1	ref|XP_002604543.1| hypothetical protein BRAFLDRAFT_265189 [B. floridae]	4.00E-35	26520| similar to restin
214	3-26_TriplEx25	1	ref|XP_002588821.1| hypothetical protein BRAFLDRAFT_89748 [B. floridae]	4.00E-35	109980| MGC84139 protein
215	Contig162	2	gb|AAN52391.1| ribosomal protein S25 [B. belcheri]	5.00E-35	61377| 40S ribosome protein S25
216	34-37_TriplEx25	1	ref|XP_002119561.1| PREDICTED: similar to IQ motif containing K [C. intestinalis]	5.00E-35	93604| hyothetical protein
217	Contig93	2	ref|XP_002585726.1| hypothetical protein BRAFLDRAFT_277639 [B. floridae]	6.00E-35	277651| meiosis expressed gene 1
218	Contig55	2	ref|XP_002604908.1| hypothetical protein BRAFLDRAFT_121622 [B. floridae]	8.00E-35	121622| similar to keratin associated protein9.3
219	29-40_TriplEx25	1	ref|XP_002587523.1| hypothetical protein BRAFLDRAFT_284483 [B. floridae]	1.00E-34	284483| splicing factor 3b, subunit 5, 10kDa
220	3-62_TriplEx25	1	ref|XP_002604955.1| hypothetical protein BRAFLDRAFT_126693 [B. floridae	4.00E-34	126693| cathepsin Z
221	34-23_TriplEx25	1	ref|XP_002601232.1| hypothetical protein BRAFLDRAFT_95011 [B. floridae]	4.00E-34	95011| hypothetical protein
222	Contig107	3	ref|XP_002604722.1| hypothetical protein BRAFLDRAFT_122567 [B. floridae]	5.00E-34	126207| elongation factor 1-alpha
223	9-4_TriplEx25	1	gb|ACO51960.1| NADH dehydrogenase iron-sulfur protein 7 [Rana catesbeiana]	5.00E-34	99644| NADH ubiquinone oxidoreductase 20kDa subunit
224	Contig106	2	ref|XP_002601285.1| hypothetical protein BRAFLDRAFT_265934 [B. floridae]	1.00E-33	265934| cytoshrome c oxydase polypeptide Vic
225	42-30_TriplEx25	1	ref|XP_002598555.1| hypothetical protein BRAFLDRAFT_66946 [B. floridae]	1.00E-33	66946| similar to Slit1-prov protein
226	29-51_TriplEx25	1	ref|XP_002587110.1| hypothetical protein BRAFLDRAFT_102202 [B. floridae]	1.00E-33	100202| growth hormone-inducible transmembrane protein
227	36-56_TriplEx25	1	ref|XP_002601974.1| hypothetical protein BRAFLDRAFT_98927 [B. floridae]	2.00E-33	5703| ATP-citrate synthase
228	Contig31	3	ref|XP_002594490.1| hypothetical protein BRAFLDRAFT_124962 [B. floridae]	2.00E-33	124962| melanotransferrin
229	27-29_TriplEx25	1	ref|XP_001190384.1| PREDICTED: similar to endonuclease[S. purpuratus]	2.00E-33	92265| similar to predicted CDS, reverse transcriptase family member
230	7-39_TriplEx25	1	ref|XP_002613003.1| hypothetical protein BRAFLDRAFT_213228 [B. floridae]	3.00E-33	223269| soc-2 suppressor of clear homolog
231	Contig124	17	gb|AAT39418.1| rhamnose-binding lectin precursor [B. belcheri]	3.00E-33	110271| rhamnose-binding lectin precursor
232	Contig118	2	ref|XP_002591613.1| hypothetical protein BRAFLDRAFT_122660 [B. floridae]	4.00E-33	122660| Gava(A) receptor associated protein
233	23-63_TriplEx25	1	ref|XP_002595369.1| hypothetical protein BRAFLDRAFT_113856 [B. floridae]	1.00E-32	113856| similar to splicing factor U2AF homolog
234	26-39_TriplEx25	1	ref|XP_002612478.1| hypothetical protein BRAFLDRAFT_75408 [B. floridae]	2.00E-32	114451| translocon-associated protein alpha subunit precursor
235	Contig66	2	gb|AAT39418.1| rhamnose-binding lectin precursor [B. belcheri]	2.00E-32	110271| rhamnose-binding lectin precursor
236	4-9_TriplEx25	1	ref|XP_002610392.1| hypothetical protein BRAFLDRAFT_209284 [B. floridae]	3.00E-32	125343| MGC68589 protein
237	Contig102	3	gb|AAN86980.1| ribosomal protein S27 [B. belcher]	4.00E-32	131078| 40S ribosomal protein S27
238	13-79_TriplEx25	1	gb|AAQ83889.1| clathrin-associated adaptor complex AP-1 small chain [B. belcheri]	5.00E-32	278617| clathrin-associated adaptor complex AP-1 small chain sigma1
239	20-47_TriplEx25	1	ref|XP_002608270.1| hypothetical protein BRAFLDRAFT_125081 [B. floridae]	1.00E-31	125081| G protein-coupled receptor kinase 4
240	2-51_TriplEx25	1	ref|XP_002587145.1| hypothetical protein BRAFLDRAFT_241911 [B. floridae]	1.00E-31	241911| similar to mitochondrial ribosomal protein S11 isoform 1
241	42-72_TriplEx25	1	gb|AAQ21039.1| ferritin [B. belcheri]	1.00E-31	288000| ferritin
242	48-45_TriplEx25	1	ref|XP_002603931.1| hypothetical protein BRAFLDRAFT_131258 [B. floridae]	2.00E-31	131258| similar to basic transcription factor 3 isoform 2
243	26-5_TriplEx25	1	ref|XP_002600258.1| hypothetical protein BRAFLDRAFT_276428 [B. floridae]	2.00E-31	276428| DC2-related axonemal dynein intermediate chain 4
244	33-16_TriplEx25	1	ref|XP_001196407.1| PREDICTED: similar to endonuclease [S. purpuratus]	2.00E-31	No hit
245	5-42_TriplEx25	1	ref|XP_002592007.1| hypothetical protein BRAFLDRAFT_79596 [B. floridae]	3.00E-31	104908| CG18471-PA
246	17-64_TriplEx25	1	ref|XP_002592165.1| hypothetical protein BRAFLDRAFT_88112 [B. floridae]	8.00E-31	85041| similar to transient receptor potential cation channel
247	49-16_TriplEx25	1	ref|XP_002614019.1| hypothetical protein BRAFLDRAFT_67397 [B. floridae]	1.00E-30	67397| hypothetical protein
248	2-83_TriplEx25	1	ref|XP_002611153.1| hypothetical protein BRAFLDRAFT_88448 [B. floridae]	1.00E-30	106757| kelch-like 9
249	16-27_TriplEx25	1	ref|XP_002613295.1| hypothetical protein BRAFLDRAFT_118710 [B. floridae]	5.00E-30	284936| hypothetical protein
250	Contig68	3	ref|XP_002606094.1| hypothetical protein BRAFLDRAFT_125111 [B. floridae]	1.00E-29	126527| ATP synthase D chain, mitochondrial
251	26-36_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	2.00E-29	68302| similar to Hyaluronidase 1
252	38-76_TriplEx25	1	dbj|BAH86195.1| NADH dehydrogenase subunit 6 [B. belcheri]	2.00E-29	No hit
253	9-73_TriplEx25	1	ref|XP_002612558.1| hypothetical protein BRAFLDRAFT_219549 [B. floridae]	3.00E-29	219549| cytochrome c oxidase
254	27-30_TriplEx25	1	ref|XP_002593030.1| hypothetical protein BRAFLDRAFT_120696 [B. floridae]	7.00E-29	120696| helical cytokine receptor CRFB4
255	6-96_TriplEx25	1	ref|XP_975635.1| similar to translation elongation factor 2 [Tribolium castaneum]	1.00E-28	281567| elongation factor 2
256	16-58_TriplEx25	1	ref|XP_002592564.1| hypothetical protein BRAFLDRAFT_68886 [B. floridae]	1.00E-28	68886| hyothetical protein
257	40-59_TriplEx25	1	ref|XP_002612910.1| hypothetical protein BRAFLDRAFT_115527 [B. floridae]	2.00E-28	286563| vacuolar ATP synthase subunit F
258	Contig125	2	ref|XP_002609016.1| hypothetical protein BRAFLDRAFT_114849 [B. floridae]	2.00E-28	116624| MGC84715 protein
259	46-37_TriplEx25	1	ref|XP_002593360.1| hypothetical protein BRAFLDRAFT_277105 [B. floridae]	2.00E-28	277105| similar to NADH dehydrogen
260	1-35_TriplEx25	1	ref|XP_002602183.1| hypothetical protein BRAFLDRAFT_76871 [B. floridae]	3.00E-28	76871| peptideglycan recognition protein S1a
261	10-5_TriplEx25	1	ref|XP_002593069.1| hypothetical protein BRAFLDRAFT_277920 [B. floridae]	3.00E-28	277920| bystin
262	Contig92	2	gb|AAT39880.1| cytochrome c oxidase subunit VIIc [B. belcheri]	4.00E-28	278223| cytochrome c oxydase subunit VIIc
263	17-4_TriplEx25	1	gb|ACY72387.1| selenium-dependent glutathione peroxidase [Hyriopsis cumingii]	2.00E-27	131284| glutathione peroxidase 1
264	4-54_TriplEx25	1	gb|AAT39880.1| cytochrome c oxidase subunit VIIc [B. belcheri]	2.00E-27	278223| cytochrome c oxydase subunit VIIc
265	40-53_TriplEx25	1	gb|AAT39880.1| cytochrome c oxidase subunit VIIc [B. belcheri]	3.00E-27	278223| cytochrome c oxidase subunit VIIc
266	5-62_TriplEx25	1	gb|AAN52372.1| ribosomal protein P2 [B. belcheri]	3.00E-27	129412| 60S acidic ribosomal protein P2
267	Contig99	3	gb|AAP21827.1| ribosomal protein S29 [B. belcheri]	4.00E-27	224271| ribosomal protein S29
268	33-35_TriplEx25	1	ref|XP_002608161.1| hypothetical protein BRAFLDRAFT_125872 [B. floridae]	6.00E-27	125872| lipoyltransferase
269	27-33_TriplEx25	1	ref|XP_002591385.1| hypothetical protein BRAFLDRAFT_86892 [B. floridae]	6.00E-27	71754| tripartite motif protein L-TRIM
270	37-80_TriplEx25	1	ref|XP_002611872.1| hypothetical protein BRAFLDRAFT_123345 [B. floridae]	2.00E-26	108728| hypothetical protein
271	41-33_TriplEx25	1	ref|XP_002602739.1| hypothetical protein BRAFLDRAFT_97705 [B. floridae]	2.00E-26	120020| similar to kelch-like protein KLHL6
272	6-68_TriplEx25	1	ref|XP_002593089.1| hypothetical protein BRAFLDRAFT_120189 [B. floridae]	2.00E-26	120716| /hypothetical protein
273	27-16_TriplEx25	1	ref|XP_002608242.1| hypothetical protein BRAFLDRAFT_125057 [B. floridae]	3.00E-26	125057| mucin-5B precursor
274	Contig13	16	ref|XP_002593089.1| hypothetical protein BRAFLDRAFT_120189 [B. floridae]	3.00E-26	120716| hypothetical protein
275	Contig59	15	ref|XP_002593089.1| hypothetical protein BRAFLDRAFT_120189 [B. floridae]	4.00E-26	120716| hypothetical protein
276	17-48_TriplEx25	1	ref|XP_002600073.1| hypothetical protein BRAFLDRAFT_58807 [B. floridae]	6.00E-26	114951| MGC83128 protein
277	3-95_TriplEx25	1	ref|XP_002608566.1| hypothetical protein BRAFLDRAFT_236042 [B. floridae]	7.00E-26	236042| unnamed protein product
278	19-70_TriplEx25	1	ref|XP_002604581.1| hypothetical protein BRAFLDRAFT_92803 [B. floridae]	1.00E-25	92803| apextrin
279	46-11_TriplEx25	1	ref|XP_002589113.1| hypothetical protein BRAFLDRAFT_75094 [B. floridae]	2.00E-25	75094| vacuolar protein sorting 13B isoform 1
280	38-14_TriplEx25	1	gb|AAL09707.1|AF420432_1 ribosomal protein L30 [B. belcheri]	3.00E-25	60646| 60S ribosomal protein L30
281	4-87_TriplEx25	1	ref|XP_002598668.1| hypothetical protein BRAFLDRAFT_67069 [B. floridae]	1.00E-24	67075| inhibition of apoptosis protein
282	26-53_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	1.00E-24	No hit
283	Contig79	3	ref|XP_002609380.1| hypothetical protein BRAFLDRAFT_124610 [B. floridae]	2.00E-24	128859| similar to collagen, type XXII, alpha 1
284	1-39_TriplEx25	1	ref|XP_002609950.1| hypothetical protein BRAFLDRAFT_85905 [B. floridae]	3.00E-24	85905| RAT keratin, type I cytoskeletal 14
285	27-76_TriplEx25	1	ref|XP_002593089.1| hypothetical protein BRAFLDRAFT_120189 [B. floridae]	5.00E-24	120716| hypothetical protein XP_897516
286	Contig151	2	dbj|BAB97379.1| amyloid protein A [B. belcheri]	6.00E-24	246351| amyloid protein A
287	46-84_TriplEx25	1	ref|XP_002594690.1| hypothetical protein BRAFLDRAFT_101450 [B. floridae]	1.00E-23	104832| hypothetical protein
288	2-25_TriplEx25	1	ref|XP_002591871.1| hypothetical protein BRAFLDRAFT_125518 [B. floridae]	1.00E-23	125518| similar to ENSANGP00000021627
289	15-35_TriplEx25	1	ref|XP_002611888.1| hypothetical protein BRAFLDRAFT_290854 [B. floridae]	2.00E-23	114730| aldehyde dehydrogenase
290	19-59_TriplEx25	1	ref|XP_002601442.1| hypothetical protein BRAFLDRAFT_286842 [B. floridae]	3.00E-23	286842| peptideglycan recongition protein
291	15-39_TriplEx25	1	ref|XP_002602402.1| hypothetical protein BRAFLDRAFT_117036 [B. floridae]	6.00E-23	266702| Zinc finger protein 706
292	11-91_TriplEx25	1	ref|XP_002594490.1| hypothetical protein BRAFLDRAFT_124962 [B. floridae]	6.00E-23	124962| similar to melanotransferrin precursor
293	25-90_TriplEx25	1	ref|NP_001133838.1| Tumor protein D53 homolog [Salmo salar]	9.00E-23	No hit
294	4-57_TriplEx25	1	ref|XP_002600576.1| hypothetical protein BRAFLDRAFT_205428 [B. floridae]	2.00E-22	205428| ubiquitinol-cytochrome c reductase complex protein
295	36-96_TriplEx25	1	ref|XP_002595624.1| hypothetical protein BRAFLDRAFT_152423 [B. floridae]	2.00E-22	152423| similar to CG30373-PA
296	5-94_TriplEx25	1	ref|XP_002593089.1| hypothetical protein BRAFLDRAFT_120189 [B. floridae]	2.00E-22	120716| hypothetical protein
297	17-44_TriplEx25	1	ref|XP_002612485.1| hypothetical protein BRAFLDRAFT_120995 [B. floridae]	5.00E-22	120995| thioester-containing protein (alpha-2 macrogloblin)
298	12-66_TriplEx25	1	ref|XP_002601442.1| hypothetical protein BRAFLDRAFT_286842 [B. floridae]	5.00E-22	286842| prptideglycan recongition protein
299	1-61_TriplEx25	1	ref|XP_002600801.1| hypothetical protein BRAFLDRAFT_95091 [B. floridae]	5.00E-22	95091| hypothetical protein
300	40-68_TriplEx25	1	ref|XP_002595715.1| hypothetical protein BRAFLDRAFT_200544 [B. floridae]	6.00E-22	227439| acyl-carrier subunit of NADH:ubiquinone oxidoreductase

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
301	15-2_TriplEx25	1	ref|XP_002591547.1| hypothetical protein BRAFLDRAFT_126746 [B. floridae]	7.00E-22	126746 l similar to carboxylesterase
302	34-73_TriplEx25	1	ref|XP_002586562.1| hypothetical protein BRAFLDRAFT_106344 [B. floridae]	1.00E-21	106344l hypothetical protein
303	12-31_TriplEx25	1	ref|XP_002596599.1| hypothetical protein BRAFLDRAFT_219212 [B. floridae]	2.00E-21	219212l similar to crystalin, gamma S
304	48-21_TriplEx25	1	ref|XP_002595583.1| hypothetical protein BRAFLDRAFT_117504 [B. floridae]	2.00E-21	117504l glycoprotein X precursor
305	11-61_TriplEx25	1	dbj|BAH86186.1| cytochrome c oxidase subunit I [B. belcheri]	2.00E-21	No hit
306	38-32_TriplEx25	1	gb|ACY72387.1| selenium-dependent glutathione peroxidase [Hyriopsis cumingii]	3.00E-21	13128l glutathione peroxidase 1
307	46-44_TriplEx25	1	ref|XP_002599163.1| hypothetical protein BRAFLDRAFT_118863 [B. floridae]	8.00E-21	117406l histone H1-delta
308	38-19_TriplEx25	1	dbj|BAH86186.1| cytochrome c oxidase subunit I [B. belcheri]	8.00E-21	No hit
309	24-93_TriplEx25	1	ref|XP_969173.1| similar to sporozoite surface protein [Tribolium castaneum]	3.00E-20	No hit
310	3-82_TriplEx25	1	ref|XP_002598334.1| hypothetical protein BRAFLDRAFT_119180 [B. floridae]	3.00E-20	119180l glycoprotein X precursor
311	22-83_TriplEx25	1	ref|XP_002215281.1| hypothetical protein BRAFLDRAFT_121759 [B. floridae]	4.00E-20	121759l erythrocyte membrane protein band 4.1-like 2
312	47-37_TriplEx25	1	ref|XP_002609281.1| hypothetical protein BRAFLDRAFT_86808 [B. floridae]	5.00E-20	86808l RAS protein activator like 2 isoform 7
313	Contig170	7	sp|P91754.1|ACT_LUMRU RecName: Full=Actin	8.00E-20	191354l actin
314	4-5_TriplEx25	1	ref|XP_002596221.1| hypothetical protein BRAFLDRAFT_260689 [B. floridae]	1.00E-19	260689l lipopolysaccharide-induced tumor necrosis factor-alpha factor
315	Contig49	2	ref|XP_002594374.1| hypothetical protein BRAFLDRAFT_72217 [B. floridae]	1.00E-19	130710l similar to leucine-rich repeat kinase2
316	Contig81	3	gb|AAN52384.1| ribosomal protein P1 [B. belcheri]	1.00E-19	240260l ribosomal protein, acidic protein 1
317	34-36_TriplEx25	1	ref|XP_002592146.1| hypothetical protein BRAFLDRAFT_124073 [B. floridae]	2.00E-19	130857l PHD finger protein
318	45-85_TriplEx25	1	gb|AAQ18145.1| p8 protein [B. belcheri]	2.00E-19	No hit
319	7-34_TriplEx25	1	ref|XP_002601998.1| hypothetical protein BRAFLDRAFT_82584 [B. floridae]	4.00E-19	65286l secreted frizzled-related protein
320	25-56_TriplEx25	1	ref|XP_002601111.1| hypothetical protein BRAFLDRAFT_75560 [B. floridae]	1.00E-18	75560l hypothetical protein
321	41-73_TriplEx25	1	ref|XP_002607332.1| hypothetical protein BRAFLDRAFT_119209 [B. floridae]	2.00E-18	119209l hypothetical protein
322	30-70_TriplEx25	1	ref|XP_002605036.1| hypothetical protein BRAFLDRAFT_124132 [B. floridae]	2.00E-18	92515l nucleotide binding protein 1
323	30-64_TriplEx25	1	ref|XP_002601983.1| hypothetical protein BRAFLDRAFT_98933 [B. floridae]	2.00E-18	98933l similar to polycystin 1-like 2
324	35-75_TriplEx25	1	ref|XP_002610164.1| hypothetical protein BRAFLDRAFT_216949 [B. floridae]	4.00E-18	216949l Acyl-CoA synthetases
325	Contig105	12	ref|XP_001892989.1| hypothetical protein Bm1_07595 [Brugia malayi]	5.00E-18	77678l hypothetical protein
326	14-63_TriplEx25	1	ref|XP_002602971.1| hypothetical protein BRAFLDRAFT_248283 [B. floridae]	6.00E-18	226736l similar to betaine homocysteine methyl transferase
327	45-83_TriplEx25	1	dbj|BAB97379.1| amyloid protein A [B. belcheri]	6.00E-18	246351l amyloid protein A
328	11-43_TriplEx25	1	ref|XP_002607045.1| hypothetical protein BRAFLDRAFT_57357 [B. floridae]	2.00E-17	226885l similar to WD repeat domain protein
329	30-55_TriplEx25	1	ref|XP_002592146.1| hypothetical protein BRAFLDRAFT_124073 [B. floridae]	2.00E-17	130857l Zn-finger-like, PHD finger
330	44-48_TriplEx25	1	gb|AAT45380.1| apextrin [B. belcheri]	2.00E-17	126750l hypothetical protein
331	35-16_TriplEx25	1	ref|XP_002588271.1| hypothetical protein BRAFLDRAFT_59692 [B. floridae]	4.00E-17	59692l ABC transporter
332	12-61_TriplEx25	1	ref|XP_002591832.1| hypothetical protein BRAFLDRAFT_125327 [B. floridae]	5.00E-17	125327l similar to UDP-N-acteylglucosamine pyrophosphorylase
333	15-91_TriplEx25	1	ref|XP_974620.1| similar to CG14235 CG14235-PA [Tribolium castaneum]	7.00E-17	No hit
334	22-3_TriplEx25	1	gb|AAA75561.1| LacOPZ-alpha peptide from pUC9 [unidentified cloning vector]	7.00E-17	No hit
335	45-87_TriplEx25	1	ref|XP_002594138.1| hypothetical protein BRAFLDRAFT_211513 [B. floridae]	8.00E-17	211513l hypothetical protein
336	25-91_TriplEx25	1	ref|XP_002597823.1| hypothetical protein BRAFLDRAFT_130184 [B. floridae]	9.00E-17	285015l hypothetical protein
337	46-4_TriplEx25	1	gb|AAZ66135.1| ribosomal protein L29 [Holosticha sp. WJC-2003]	9.00E-17	276720l ribosomal L29e protein
338	19-37_TriplEx25	1	ref|XP_002613686.1| hypothetical protein BRAFLDRAFT_287967 [B. floridae]	1.00E-16	112583l alpha-mannosidase
339	Contig147	2	ref|XP_002597718.1| hypothetical protein BRAFLDRAFT_114379 [B. floridae]	1.00E-16	114379l hypothetical protein
340	Contig8	2	ref|XP_002595984.1| hypothetical protein BRAFLDRAFT_96757 [B. floridae]	1.00E-16	96757l hypothetical protein
341	Contig75	2	ref|XP_002611841.1| hypothetical protein BRAFLDRAFT_83135 [B. floridae]	2.00E-16	83135l 240 kDa protein of rod photoreceptor CNG-channel
342	10-67_TriplEx25	1	ref|XP_002597628.1| hypothetical protein BRAFLDRAFT_225970 [B. floridae]	2.00E-16	225970l cytochrome P450 family 2
343	10-7_TriplEx25	1	gb|ABU39931.1| beta-galactosidase [Cloning vector pGreenII 0179]	2.00E-16	No hit
344	15-15_TriplEx25	1	ref|XP_002606423.1| hypothetical protein BRAFLDRAFT_67675 [B. floridae]	3.00E-16	67675l L-iditol 2-dehydrogenase
345	Contig143	11	ref|XP_002596298.1| hypothetical protein BRAFLDRAFT_82099 [B. floridae]	3.00E-16	112027l putative cyclic nucleotide gated channel beta
346	42-49_TriplEx25	1	ref|XP_001191411.1| PREDICTED: similar to cubilin [S. purpuratus]	3.00E-16	75213l hypothetical protein
347	5-38_TriplEx25	1	dbj|BAB97381.1| rab GDP-dissociation inhibitor [B. belcheri]	3.00E-16	105147l RAB GDP-dissociation inhibitor
348	19-6_TriplEx25	1	gb|AAN52382.1| ribosomal protein L39 [B. belcheri]	4.00E-16	No hit
349	47-17_TriplEx25	1	ref|XP_002598529.1| hypothetical protein BRAFLDRAFT_118317 [B. floridae]	5.00E-16	118319l ribosomal protein S26E
350	5-45_TriplEx25	1	ref|XP_002601832.1| hypothetical protein BRAFLDRAFT_75950 [B. floridae]	6.00E-16	75950l myosin VC
351	15-45_TriplEx25	1	gb|AAI66142.1| Xenotropic and polytropic retrovirus receptor [Xenopus tropicalis]	7.00E-16	No hit
352	20-31_TriplEx25	1	ref|XP_002605434.1| hypothetical protein BRAFLDRAFT_74249 [B. floridae]	1.00E-15	84153l conserved hypothetical protein
353	Contig96	2	ref|XP_002131762.1| PREDICTED: hypothetical protein isoform 1 [C. intestinalis]	1.00E-15	No hit
354	13-57_TriplEx25	1	ref|XP_002605381.1| hypothetical protein BRAFLDRAFT_74201 [B. floridae]	3.00E-15	74201l hypothetical protein
355	38-36_TriplEx25	1	ref|XP_002590850.1| hypothetical protein BRAFLDRAFT_125712 [B. floridae]	3.00E-15	62639l mechanosensory abnormality protein 2, isoform b
356	4-65_TriplEx25	1	ref|XP_002612669.1| hypothetical protein BRAFLDRAFT_78700 [B. floridae]	6.00E-15	86714l similar to vesicular inhibitory amino acid transporter
357	34-54_TriplEx25	1	ref|XP_001650445.1| hypothetical protein AaeL_AAEL015064 [Aedes aegypti]	7.00E-15	253759l SMT3 suppressor of mif two 3 homolog 1
358	20-39_TriplEx25	1	ref|XP_002606533.1| hypothetical protein BRAFLDRAFT_270557 [B. floridae]	1.00E-14	131075l similar to CG13855-PA
359	17-86_TriplEx25	1	ref|XP_002591726.1| hypothetical protein BRAFLDRAFT_80820 [B. floridae]	1.00E-14	80820l versican V2 splice-variant precursor
360	17-57_TriplEx25	1	ref|XP_002587955.1| hypothetical protein BRAFLDRAFT_87343 [B. floridae]	1.00E-14	201861l extracellular peptidase inhibitor
361	1-31_TriplEx25	1	ref|XP_002604955.1| hypothetical protein BRAFLDRAFT_126693 [B. floridae]	2.00E-14	126693l cathepsin Z
362	40-34_TriplEx25	1	ref|XP_002594720.1| hypothetical protein BRAFLDRAFT_81173 [B. floridae]	6.00E-14	88316l similar to tenascin R
363	Contig29	5	ref|XP_002599368.1| hypothetical protein BRAFLDRAFT_64274 [B. floridae]	9.00E-14	64274l serine threonine-protein kinase WNK2
364	19-32_TriplEx25	1	ref|XP_002609719.1| hypothetical protein BRAFLDRAFT_102466 [B. floridae]	1.00E-13	102466l X-linked interleukin-1 receptor accessory protein-like 1 precursor
365	38-24_TriplEx25	1	ref|XP_002604792.1| hypothetical protein BRAFLDRAFT_119478 [B. floridae	1.00E-13	126726l hypothetical protein
366	29-70_TriplEx25	1	ref|XP_002604734.1| hypothetical protein BRAFLDRAFT_265555 [B. floridae]	1.00E-13	265555l DEAH (Asp-Glu-Ala-His) box polypeptide 37, helicase
367	35-13_TriplEx25	1	ref|XP_002611493.1| hypothetical protein BRAFLDRAFT_117197 [B. floridae]	2.00E-13	117197l hypothetical protein
368	30-27_TriplEx25	1	ref|XP_002608792.1| hypothetical protein BRAFLDRAFT_125597 [B. floridae]	2.00E-13	91200l mitochondrial ATP synthase coupling factor 6
369	47-50_TriplEx25	1	ref|XP_001656348.1| copper chaperone Atox1, putative [Aedes aegypti]	2.00E-13	249962l copper transport protein
370	43-72_TriplEx25	1	dbj|BAF91637.1| M polyprotein [Tinaroo virus]	2.00E-13	132560l major yolk protein
371	25-38_TriplEx25	1	ref|XP_002610900.1| hypothetical protein BRAFLDRAFT_91496 [B. floridae]	3.00E-13	91496l similar to proprotein convertase 6B
372	38-33_TriplEx25	1	ref|XP_002613082.1| hypothetical protein BRAFLDRAFT_89964 [B. floridae	4.00E-13	89964l similar to putative protein family member
373	33-90_TriplEx25	1	ref|XP_002585915.1| hypothetical protein BRAFLDRAFT_110789 [B. floridae]	4.00E-13	110789l hypothetical protein
374	4-3_TriplEx25	1	ref|XP_002599157.1| hypothetical protein BRAFLDRAFT_68766 [B. floridae]	5.00E-13	68766l similar to chromosome 17 open reading frame 27
375	32-87_TriplEx25	1	ref|XP_002601555.1| hypothetical protein BRAFLDRAFT_230625 [B. floridae]	6.00E-13	230625l similar to Cezanne 2 protein
376	9-93_TriplEx25	1	ref|XP_002598730.1| hypothetical protein BRAFLDRAFT_282817 [B. floridae]	7.00E-13	282817l DNA-directed RNA polymerase III subunit 127.6kDa polypeptide
377	31-35_TriplEx25	1	ref|XP_002596302.1| hypothetical protein BRAFLDRAFT_82095 [B. floridae]	8.00E-13	82095l hypothetical protein
378	11-11_TriplEx25	1	gb|ACN10033.1| 60S ribosomal protein L38 [Salmo salar]	9.00E-13	286891l ribosomal protein L38e
379	22-56_TriplEx25	1	ref|XP_002589965.1| hypothetical protein BRAFLDRAFT_122930 [B. floridae]	1.00E-12	12449l trichohyalin
380	Contig86	2	gb|AAL09708.1|AF420433_1 ribosomal protein L38 [B. belcheri]	1.00E-12	286891l ribosomal L38e protein
381	Contig5	2	ref|XP_002596663.1| hypothetical protein BRAFLDRAFT_78440 [B. floridae]	2.00E-12	78440l hypothetical protein
382	33-93_TriplEx25	1	ref|XP_001929599.1| similar to mitochondrial ribosomal protein [Sus scrofa]	2.00E-12	133371l similar to mitochondrial ribosomal protein L14 isoform 1
383	12-21_TriplEx25	1	ref|XP_002611121.1| hypothetical protein BRAFLDRAFT_164206 [B. floridae]	3.00E-12	104677l tubulin tyrosine ligase
384	15-1_TriplEx25	1	dbj|BAD86652.1| reverse transcriptase [Bombyx mori]	3.00E-12	No hit
385	35-69_TriplEx25	1	ref|XP_002606019.1| hypothetical protein BRAFLDRAFT_269786 [B. floridae]	4.00E-12	269786l similar to NADH dehydrogenase
386	Contig154	2	ref|XP_002596719.1| hypothetical protein BRAFLDRAFT_78381 [B. floridae]	4.00E-12	78381l S-adenosylmethionine synthetase isoform type-2
387	18-13_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	5.00E-12	207197l hypothetical protein
388	27-7_TriplEx25	1	ref|XP_001062247.1| PREDICTED: hypothetical protein [Rattus norvegicus]	6.00E-12	263902l similar to synaptotagmin
389	24-47_TriplEx25	1	ref|XP_001640657.1| predicted protein [Nematostella vectensis]	9.00E-12	108630l hypothetical protein
390	4-92_TriplEx25	1	ref|XP_002596251.1| hypothetical protein BRAFLDRAFT_65994 [B. floridae]	1.00E-11	65994l glycoside hydrolase, family 2, sugar binding
391	32-65_TriplEx25	1	ref|XP_002588127.1| hypothetical protein BRAFLDRAFT_124953 [B. floridae]	1.00E-11	124953l hypothetical protein
392	47-2_TriplEx25	1	ref|XP_002586342.1| hypothetical protein BRAFLDRAFT_108822 [B. floridae]	2.00E-11	108822l epicuticlin
393	7-63_TriplEx25	1	ref|XP_002611587.1| hypothetical protein BRAFLDRAFT_117157 [B. floridae]	4.00E-11	117157l hypothetical protein
394	30-46_TriplEx25	1	ref|XP_002610900.1| hypothetical protein BRAFLDRAFT_91496 [B. floridae]	8.00E-11	91496l similar to proprotein convertase 6B
395	4-30_TriplEx25	1	ref|XP_002596223.1| hypothetical protein BRAFLDRAFT_202938 [B. floridae]	1.00E-10	256329l AAAdomain containing 1b
396	17-76_TriplEx25	1	ref|XP_002609805.1| hypothetical protein BRAFLDRAFT_219461 [B. floridae]	2.00E-10	219461l similar to SMC5 protein isoform 1
397	Contig97	2	ref|XP_002607419.1| hypothetical protein BRAFLDRAFT_205068 [B. floridae]	2.00E-10	205068l MGC69156 protein
398	43-87_TriplEx25	1	ref|XP_002603178.1| hypothetical protein BRAFLDRAFT_93418 [B. floridae]	2.00E-10	93418l glanylate cyclase
399	Contig73	2	ref|XP_002588187.1| hypothetical protein BRAFLDRAFT_113823 [B. floridae]	2.00E-10	113823l ribosomal protein L5
400	37-79_TriplEx25	1	ref|XP_001178247.1| PREDICTED: hypothetical protein [S. purpuratus]	2.00E-10	No hit

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
401	2-36_TriplEx25	1	ref|XP_002593326.1| hypothetical protein BRAFLDRAFT_70894 [B. floridae]	3.00E-10	83874| similar to Dedicator of cytokinesis protein 9
402	22-95_TriplEx25	1	ref|NP_852472.1| notch-regulated ankyrin repeat protein a [Danio rerio]	4.00E-10	62342| unknown
403	21-96_TriplEx25	1	gb|AAG34526.1| beta-galactosidase [Cloning vector pUG26]	4.00E-10	No hit
404	15-21_TriplEx25	1	ref|XP_002201658.1| hypothetical protein BRAFLDRAFT_56868 [B. floridae]	5.00E-10	246042| rab-like protein
405	20-38_TriplEx25	1	ref|XP_002589093.1| hypothetical protein BRAFLDRAFT_75073 [B. floridae]	6.00E-10	89210| hypothetical protein
406	Contig165	2	ref|XP_002122671.1| Lipoxygenase homology domain-containing protein [C. intestinalis]	7.00E-10	125536| similar to hypothetical protein
407	Contig18	2	ref|XP_001178872.1| PREDICTED: hypothetical protein [S. purpuratus]	7.00E-10	No hit
408	14-52_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	7.00E-10	No hit
409	17-19_TriplEx25	1	ref|XP_002613710.1| hypothetical protein BRAFLDRAFT_130686 [B. floridae]	2.00E-09	130686| glutamate dehydrogenase 1a
410	16-57_TriplEx25	1	ref|XP_002603498.1| hypothetical protein BRAFLDRAFT_220050 [B. floridae]	2.00E-09	220050| similar to diacylglycerol kinase, delta 130kDa isoform 1
411	7-35_TriplEx25	1	ref|XP_002597882.1| hypothetical protein BRAFLDRAFT_234174 [B. floridae]	2.00E-09	109876| hypothetical protein
412	40-94_TriplEx25	1	ref|XP_002604428.1| hypothetical protein BRAFLDRAFT_58757 [B. floridae]	3.00E-09	62099| dual specificity protein phosphatase
413	48-77_TriplEx25	1	gb|AAQ57129.1| endonuclease and reverse transcriptase-like protein [Bombyx mori]	3.00E-09	No hit
414	5-23_TriplEx25	1	ref|XP_002594741.1| hypothetical protein BRAFLDRAFT_81191 [B. floridae]	4.00E-09	81207| similar to BRICHOS doomain containing protein
415	8-61_TriplEx25	1	ref|XP_001508257.1| similar to Kinesin-like protein [Ornithorhynchus anatinus]	4.00E-09	117050| kinesin-like protein
416	29-30_TriplEx25	1	gb|EDL90354.1| Fas-associated factor 1 [Rattus norvegicus]	4.00E-09	284504| Fas (TNFRSF6) associated factor 1
417	5-11_TriplEx25	1	ref|XP_002596298.1| hypothetical protein BRAFLDRAFT_82099 [B. floridae]	5.00E-09	11202| putative cyclic nucleotide gated channel beta 1
418	25-53_TriplEx25	1	ref|XP_002592636.1| hypothetical protein BRAFLDRAFT_85084 [B. floridae]	5.00E-09	85084| similar to 28kD interferon responsive protein
419	27-11_TriplEx25	1	ref|XP_002609202.1| hypothetical protein BRAFLDRAFT_125959 [B. floridae]	8.00E-09	128913| similar to nerve growth factor receptor
420	2-77_TriplEx25	1	ref|XP_002594429.1| hypothetical protein BRAFLDRAFT_72185 [B. floridae]	8.00E-09	205401| CG17906-PA
421	8-18_TriplEx25	1	ref|XP_002610246.1| hypothetical protein BRAFLDRAFT_92966 [B. floridae]	1.00E-08	257638| hypothetical protein
422	2-74_TriplEx25	1	ref|XP_002604175.1| hypothetical protein BRAFLDRAFT_278184 [B. floridae]	1.00E-08	278184| Sac domain-containing inositol phosphatase 3
423	35-84_TriplEx25	1	ref|XP_002599845.1| hypothetical protein BRAFLDRAFT_230181 [B. floridae]	1.00E-08	230181| lecithin:cholesterol acyltransferase
424	1-81_TriplEx25	1	ref|XP_002071607.1| GK10072 [Drosophila willistoni]	1.00E-08	126283| 26S protease regulatory subunit 7
425	7-7_TriplEx25	1	ref|XP_001204249.1| PREDICTED: similar to C18orf34 protein [S. purpuratus]	2.00E-08	No hit
426	Contig163	2	ref|XP_002599451.1| hypothetical protein BRAFLDRAFT_223882 [B. floridae]	3.00E-08	112648| hypothetical protein
427	32-42_TriplEx25	1	ref|XP_002591097.1| hypothetical protein BRAFLDRAFT_108708 [B. floridae]	3.00E-08	109614| variable lymphocyte receptor A
428	17-78_TriplEx25	1	dbj|BAC82626.1| pol-like protein [C. intestinalis]	4.00E-08	223959| similar to kinesin-like protein
429	17-18_TriplEx25	1	ref|XP_395748.2| PREDICTED: similar to FK506-binding protein [Apis mellifera]	5.00E-08	288367| FK506 binding protein 4
430	46-92_TriplEx25	1	ref|XP_002604765.1| hypothetical protein BRAFLDRAFT_119473 [B. floridae]	5.00E-08	119473| hypothetical Zinc-containing alcohol dehydrogenase superfamily
431	15-14_TriplEx25	1	ref|XP_002611794.1| hypothetical protein BRAFLDRAFT_99048 [B. floridae]	6.00E-08	109477| similar to metabotropic glutamate receptor precursor
432	32-57_TriplEx25	1	ref|XP_002608162.1| hypothetical protein BRAFLDRAFT_125871 [B. floridae]	1.00E-07	125871| dermatopontin
433	7-22_TriplEx25	1	ref|XP_002613680.1| hypothetical protein BRAFLDRAFT_250332 [B. floridae]	2.00E-07	250332| lysosomal cofactor/neurotrophic factor prosaposin
434	25-68_TriplEx25	1	ref|XP_002609241.1| hypothetical protein BRAFLDRAFT_90695 [B. floridae]	2.00E-07	73989| TPR repeat:Tetratricopeptide TPR_3
435	46-48_TriplEx25	1	ref|XP_002592500.1| hypothetical protein BRAFLDRAFT_68991 [B. floridae]	2.00E-07	94779| hypothetical protein
436	35-11_TriplEx25	1	ref|XP_002590692.1| hypothetical protein BRAFLDRAFT_89494 [B. floridae]	2.00E-07	78109| hypothetical protein
437	17-60_TriplEx25	1	gb|AAF26301.1|AF184616_1 proprotein convertase aPC6B isoform [B. californiense]	3.00E-07	85770| proprotein convertase subtilisin/kexin type 5 precursor
438	42-60_TriplEx25	1	ref|XP_002588157.1| hypothetical protein BRAFLDRAFT_68795 [B. floridae]	3.00E-07	68795| erythrocyte membrane protein 1
439	Contig21	2	ref|XP_002593165.1| hypothetical protein BRAFLDRAFT_72756 [B. floridae]	4.00E-07	72756| similar to dynein
440	16-12_TriplEx25	1	ref|XP_002607276.1| hypothetical protein BRAFLDRAFT_88224 [B. floridae]	6.00E-07	88224| hypothetical protein
441	42-18_TriplEx25	1	ref|XP_002609936.1| hypothetical protein BRAFLDRAFT_85884 [B. floridae]	7.00E-07	85884| similar to leucine rich repeat containing 6
442	35-27_TriplEx25	1	ref|XP_002591684.1| hypothetical protein BRAFLDRAFT_223577 [B. floridae]	7.00E-07	80786| similar to calmodulin binding
443	26-41_TriplEx25	1	ref|XP_002603255.1| hypothetical protein BRAFLDRAFT_115457 [B. floridae]	9.00E-07	285231| hypothetical protein
444	36-38_TriplEx25	1	ref|XP_001180669.1| similar to endonuclease-reverse transcriptase [S. purpuratus]	9.00E-07	90613| endonuclease-reverse transcriptase
445	25-24_TriplEx25	1	ref|XP_002609625.1| hypothetical protein BRAFLDRAFT_125037 [B. floridae]	2.00E-06	125037| similar to tudor domain containing 1
446	16-17_TriplEx25	1	ref|XP_002596302.1| hypothetical protein BRAFLDRAFT_82095 [B. floridae]	2.00E-06	82095| hypothetical protein
447	25-62_TriplEx25	1	ref|XP_002131454.1| similar to Ferric-chelate reductase 1[C. intestinalis]	2.00E-06	82772| hypothetical protein
448	46-77_TriplEx25	1	ref|XP_002601388.1| hypothetical protein BRAFLDRAFT_130399 [B. floridae]	3.00E-06	123218| transcription factor AP-1
449	33-68_TriplEx25	1	ref|XP_002596301.1| hypothetical protein BRAFLDRAFT_82096 [B. floridae]	3.00E-06	112024| PLAFV S-antigen protein precursor
450	45-70_TriplEx25	1	ref|XP_002596142.1| hypothetical protein BRAFLDRAFT_66129 [B. floridae	3.00E-06	109590| putative elongation factor 2 kinase
451	Contig116	7	emb|CAM36311.1| hypothetical protein [Thermobia domestica]	3.00E-06	No hit
452	49-6_TriplEx25	1	ref|XP_002587699.1| hypothetical protein BRAFLDRAFT_94602 [B. floridae]	4.00E-06	94602| hypothetical protein
453	15-83_TriplEx25	1	ref|XP_002593002.1| hypothetical protein BRAFLDRAFT_117784 [B. floridae]	5.00E-06	74318| hypothetical protein
454	1-86_TriplEx25	1	ref|XP_002591884.1| hypothetical protein BRAFLDRAFT_89401 [B. floridae]	5.00E-06	99999| hypothetical protein
455	16-37_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	5.00E-06	No hit
456	Contig46	28	pir||JC1348 hypothetical 18K protein - goldfish mitochondrion	6.00E-06	No hit
457	24-23_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	6.00E-06	No hit
458	8-54_TriplEx25	1	ref|XP_002163289.1| similar to Leishmanolysin family protein [Hydra magnipapillata]	8.00E-06	76831| similar to riddle like (86.6kDa)
459	13-11_TriplEx25	1	dbj|BAB97379.1| amyloid protein A [B. belcheri]	8.00E-06	No hit
460	Contig109	28	emb|CAM36311.1| hypothetical protein [Thermobia domestica]	9.00E-06	No hit
461	36-11_TriplEx25	1	ref|XP_002598430.1| hypothetical protein BRAFLDRAFT_123398 [B. floridae]	1.00E-05	123398| zinc finger containing transactivation factor Sp5
462	12-65_TriplEx25	1	ref|XP_002587988.1| hypothetical protein BRAFLDRAFT_88967 [B. floridae]	1.00E-05	88967| similar to kelch-like 12 isoform 1
463	40-44_TriplEx25	1	ref|XP_001630702.1| predicted protein [Nematostella vectensis]	1.00E-05	No hit
464	1-36_TriplEx25	1	gb|AAC83651.1| beta-D-galactosidase [Integrational vector pMUTIN2]	1.00E-05	No hit
465	26-40_TriplEx25	1	ref|XP_002595676.1| hypothetical protein BRAFLDRAFT_64810 [B. floridae]	2.00E-05	64810| similar to polycystin 1; polycystic kidney disease 1
466	2-37_TriplEx25	1	ref|XP_001022111.1| hypothetical protein TTHERM_01150360 [Tetrahymena thermophila]	2.00E-05	No hit
467	Contig30	2	ref|NP_001078996.1| keratin associated protein 9-5 [Mus musculus]	2.00E-05	No hit
468	8-94_TriplEx25	1	ref|YP_002122604.1| collagen-like surface protein SclZ.5 [Streptococcus equi]	4.00E-05	No hit
469	32-56_TriplEx25	1	ref|XP_002599103.1| hypothetical protein BRAFLDRAFT_81769 [B. floridae]	4.00E-05	111628| similar to Smcr8 protein
470	32-50_TriplEx25	1	dbj|BAA82359.1| HrEpiC [Halocynthia roretzi]	5.00E-05	10594| CUB and SUSHI multiple domain
471	19-42_TriplEx25	1	ref|XP_002603369.1| hypothetical protein BRAFLDRAFT_80362 [B. floridae]	6.00E-05	71353| hypothetical kinesin
472	11-14_TriplEx25	1	gb|ABB29604.1| DNA replication licensing factor MCM7 component [Platynereis dumerilii]	7.00E-05	124874| DNA replication locensing factor MCM7
473	23-96_TriplEx25	1	ref|XP_002259201.1| hypothetical protein [Plasmodium knowlesi]	8.00E-05	128716| C5orf3 protein
474	10-42_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	9.00E-05	No hit
475	48-44_TriplEx25	1	dbj|BAB47146.1| complement component C3 [B. belcheri]	9.00E-05	No hit
476	48-43_TriplEx25	1	dbj|BAB47146.1| complement component C3 [B. belcheri]	9.00E-05	132771| complement component C3
477	37-60_TriplEx25	1	ref|ZP_06055635.1| acetate--CoA ligase [alpha proteobacterium HIMB114]	1.00E-04	No hit
478	Contig131	2	ref|XP_002613212.1| hypothetical protein BRAFLDRAFT_278055 [B. floridae]	1.00E-04	278055| heat shock protein
479	47-75_TriplEx25	1	ref|XP_002606148.1| hypothetical protein BRAFLDRAFT_126479 [B. floridae]	2.00E-04	126479| similar to nucleolar protein with MIF4G domain 1
480	44-70_TriplEx25	1	ref|XP_002134772.1| GA23662 [Drosophila pseudoobscura pseudoobscura]	2.00E-04	No hit
481	7-40_TriplEx25	1	gb|ACF33477.1| protein polymer R4 precursor [synthetic construct]	2.00E-04	131007| hypothetical protein
482	48-93_TriplEx25	1	gb|AAR19217.1| HMW glutenin subunit y [Aegilops uniaristata]	2.00E-04	131007| hypothetical protein
483	33-8_TriplEx25	1	|XP_001030500.2| Major Facilitator Superfamily protein [Tetrahymena thermophila]	2.00E-04	No hit
484	19-68_TriplEx25	1	ref|XP_002613532.1| hypothetical protein BRAFLDRAFT_71828 [B. floridae]	3.00E-04	71828| similar to plasma glutamate carboxypeptidase; aminopeptidase
485	34-85_TriplEx25	1	ref|XP_002596301.1| hypothetical protein BRAFLDRAFT_82096 [B. floridae]	3.00E-04	No hit
486	26-19_TriplEx25	1	gb|AAB64397.1| beta-galactosidase [unidentified cloning vector]	3.00E-04	No hit
487	49-68_TriplEx25	1	ref|XP_002612898.1| hypothetical protein BRAFLDRAFT_151528 [B. floridae]	4.00E-04	127279| similar to apical early endosomal glycoprotein precursor
488	8-31_TriplEx25	1	ref|XP_002604433.1| hypothetical protein BRAFLDRAFT_79267 [B. floridae]	5.00E-04	104725| tetraytricopeptide repeat protein 13
489	15-88_TriplEx25	1	ref|XP_654495.2| lysozyme [Entamoeba histolytica HM-1:IMSS]	6.00E-04	75210| Tfp pilus assembly protein FimV , contains LysM domain
490	45-37_TriplEx25	1	dbj|BAA88483.1| enolase-2 [Lethenteron reissneri]	6.00E-04	281232| alpha-enolase
491	6-50_TriplEx25	1	ref|XP_002591961.1| hypothetical protein BRAFLDRAFT_79554 [B. floridae]	9.00E-04	67812| similar to tenascin R
492	38-11_TriplEx25	1	ref|NP_001156462.1| COP9 complex homolog subunit 5 [Acyrthosiphon pisum]	9.00E-04	124400| COP9 signalosome complex subunit 5
493	18-59_TriplEx25	1	ref|XP_002601109.1| hypothetical protein BRAFLDRAFT_75558 [B. floridae]	0.001	75558| hedgehog protein
494	1-56_TriplEx25	1	ref|XP_002597226.1| hypothetical protein BRAFLDRAFT_66351 [B. floridae]	0.001	101526| ribosome-binding protein 1
495	19-5_TriplEx25	1	ref|XP_002586537.1| hypothetical protein BRAFLDRAFT_138180 [B. floridae]	0.001	138180| similar to Caspase precursor (drICE)
496	Contig76	2	gb|AAQ21039.1| ferritin [B. belcheri]	0.001	288000| ferritin
497	33-83_TriplEx25	1	gb|AAN77903.1| ferritin [B. belcheri]	0.001	No hit
498	40-63_TriplEx25	1	ref|XP_002590271.1| hypothetical protein BRAFLDRAFT_279345 [B. floridae]	0.002	279345| cystatin B
499	47-49_TriplEx25	1	gb|AAF14007.1| SUP35 homolog [Zygosaccharomyces rouxii]	0.002	No hit
500	25-43_TriplEx25	1	ref|XP_002607816.1| hypothetical protein BRAFLDRAFT_199653 [B. floridae]	0.003	199653| similar to family with sequence similarity 55, member C

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
501	30-78_TriplEx25	1	ref|XP_002588620.1| hypothetical protein BRAFLDRAFT_249937 [B. floridae]	0.003	251148| sialyltransferase 8A isoform 1
502	26-46_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	0.003	No hit
503	38-87_TriplEx25	1	ref|XP_002585585.1| hypothetical protein BRAFLDRAFT_133155 [B. floridae]	0.004	133155| zinc finger RNA binding protein
504	13-50_TriplEx25	1	gb|EDL31623.1| AT rich interactive domain 3A (Bright like), isoform CRA_b [Mus musculus]	0.004	No hit
505	9-87_TriplEx25	1	ref|XP_002613491.1| hypothetical protein BRAFLDRAFT_119839 [B. floridae]	0.005	119839| heat shock protein HSP 90-alpha
506	18-34_TriplEx25	1	ref|XP_002607537.1| hypothetical protein BRAFLDRAFT_106485 [B. floridae]	0.005	106685| hypothetical protein PY06855
507	29-10_TriplEx25	1	ref|XP_002087786.1| GE14937 [Drosophila yakuba]	0.006	125570| GA11610-PA
508	20-23_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.006	No hit
509	33-95_TriplEx25	1	ref|XP_002590886.1| hypothetical protein BRAFLDRAFT_129595 [B. floridae]	0.007	288046| similar to cytochrome c oxidase, subunit VIIa 2
510	14-35_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	0.009	No hit
511	36-15_TriplEx25	1	ref|XP_002594138.1| hypothetical protein BRAFLDRAFT_211513 [B. floridae]	0.01	258920| glutamic-oxaloacetic transaminase 1, soluble
512	Contig74	7	ref|XP_002590541.1| hypothetical protein BRAFLDRAFT_86219 [B. floridae]	0.01	No hit
513	7-58_TriplEx25	1	ref|XP_002602020.1| hypothetical protein BRAFLDRAFT_82607 [B. floridae]	0.011	82607| hypothetical protein
514	10-9_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.011	No hit
515	41-18_TriplEx25	1	ref|XP_001919308.1| PREDICTED: similar to helentron 3 [Danio rerio]	0.013	No hit
516	22-52_TriplEx25	1	ref|XP_001368603.1| PREDICTED: hypothetical protein [Monodelphis domestica]	0.013	No hit
517	31-43_TriplEx25	1	ref|XP_001095523.1| PREDICTED: similar to cell division cycle 2-like 1 [Macaca mulatta]	0.013	No hit
518	35-7_TriplEx25	1	dbj|BAD86652.1| reverse transcriptase [Bombyx mori]	0.013	239889| hypothetical protein
519	Contig139	3	gb|EFA10298.1| hypothetical protein TcasGA2_TC012513 [Tribolium castaneum]	0.014	No hit
520	7-21_TriplEx25	1	ref|XP_002606954.1| hypothetical protein BRAFLDRAFT_64950 [B. floridae]	0.015	No hit
521	33-32_TriplEx25	1	gb|AAR15424.1| Cu2+ plastocyanin-like [Sisymbrium irio]	0.015	No hit
522	32-59_TriplEx25	1	ref|XP_626647.1| hypothetical protein [Cryptosporidium parvum Iowa II]	0.017	No hit
523	37-93_TriplEx25	1	ref|XP_002592931.1| hypothetical protein BRAFLDRAFT_65518 [B. floridae]	0.017	100505| similar to THAP domain containing 4
524	5-76_TriplEx25	1	gb|EEN64517.1| hypothetical protein BRAFLDRAFT_92411 [B. floridae]	0.022	111214| similar to polycystin 1-like 2
525	34-3_TriplEx25	1	ref|XP_001962656.1| GF14327 [Drosophila ananassae]	0.025	127073| basic helix-loop-herix dimerisation region bHLH
526	15-9_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.027	No hit
527	24-11_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.028	No hit
528	8-85_TriplEx25	1	ref|XP_002605416.1| hypothetical protein BRAFLDRAFT_278505 [B. floridae]	0.033	No hit
529	15-57_TriplEx25	1	ref|XP_002605416.1| hypothetical protein BRAFLDRAFT_278505 [B. floridae]	0.034	278505| similar to proliferation-associated protein 1
530	15-51_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	0.035	No hit
531	5-67_TriplEx25	1	ref|XP_002594741.1| hypothetical protein BRAFLDRAFT_81191 [B. floridae]	0.037	81207| similar to BRICHOS domain containing protein
532	45-23_TriplEx25	1	ref|XP_002609898.1| hypothetical protein BRAFLDRAFT_125987 [B. floridae]	0.038	No hit
533	23-44_TriplEx25	1	ref|XP_002609570.1| hypothetical protein BRAFLDRAFT_129887 [B. floridae]	0.038	No hit
534	2-71_TriplEx25	1	ref|NP_001131088.1| protein tyrosine phosphatase [Xenopus tropicalis]	0.042	No hit
535	19-65_TriplEx25	1	ref|XP_799179.2| PREDICTED: similar to KIAA1370 protein [S. purpuratus]	0.044	No hit
536	38-43_TriplEx25	1	ref|XP_001178123.1| PREDICTED: similar to ORF2-encoded protein [S. purpuratus]	0.049	No hit
537	26-12_TriplEx25	1	ref|XP_002605229.1| hypothetical protein BRAFLDRAFT_126601 [B. floridae]	0.053	122694| gstromal cell derived factor 2-like protein
538	12-83_TriplEx25	1	ref|XP_001721636.2| PREDICTED: hypothetical protein [Homo sapiens]	0.053	No hit
539	33-47_TriplEx25	1	ref|XP_001631026.1| predicted protein [Nematostella vectensis]	0.055	No hit
540	19-57_TriplEx25	1	gb|AAS22104.1| small hydrophobic protein [Human metapneumovirus]	0.057	No hit
541	36-83_TriplEx25	1	ref|XP_002610515.1| hypothetical protein BRAFLDRAFT_65679 [B. floridae]	0.065	No hit
542	45-4_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.065	No hit
543	9-31_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	0.072	No hit
544	Contig6	2	ref|XP_001521089.1| PREDICTED: similar to cap binding protein [Ornithorhynchus anatinus]	0.072	No hit
545	9-32_TriplEx25	1	ref|XP_001454405.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	0.072	No hit
546	Contig57	2	ref|NP_001155106.1| transient receptor potential cation channel [Xenopus tropicalis]	0.072	No hit
547	49-79_TriplEx25	1	ref|XP_002613709.1| hypothetical protein BRAFLDRAFT_130685 [B. floridae]	0.075	No hit
548	48-71_TriplEx25	1	ref|XP_002605940.1| hypothetical protein BRAFLDRAFT_87382 [B. floridae]	0.075	No hit
549	48-95_TriplEx25	1	ref|XP_002590541.1| hypothetical protein BRAFLDRAFT_86219 [B. floridae]	0.075	No hit
550	41-72_TriplEx25	1	ref|XP_002414851.1| Ubiquitin-fold modifier-1 specific protease [Ixodes scapularis]	0.082	148252| hypothetical protein
551	29-95_TriplEx25	1	ref|XP_002601684.1| hypothetical protein BRAFLDRAFT_94561 [B. floridae]	0.085	No hit
552	35-8_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.085	No hit
553	35-4_TriplEx25	1	ref|XP_002585980.1| hypothetical protein BRAFLDRAFT_255846 [B. floridae]	0.11	No hit
554	35-61_TriplEx25	1	ref|XP_002342505.1| PREDICTED: hypothetical protein XP_002342505 [Homo sapiens]	0.11	No hit
555	44-89_TriplEx25	1	ref|XP_001452058.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	0.11	No hit
556	22-7_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.11	No hit
557	30-87_TriplEx25	1	ref|XP_002605932.1| hypothetical protein BRAFLDRAFT_124890 [B. floridae]	0.14	No hit
558	36-60_TriplEx25	1	ref|XP_002156054.1| PREDICTED: similar to predicted protein [Hydra magnipapillata]	0.14	No hit
559	6-82_TriplEx25	1	ref|XP_964717.1| hypothetical protein NCU00556 [Neurospora crassa OR74A]	0.16	No hit
560	Contig69	2	ref|XP_002596301.1| hypothetical protein BRAFLDRAFT_82096 [B. floridae]	0.16	No hit
561	4-94_TriplEx25	1	ref|XP_002408351.1| ribosomal protein L37, putative [Ixodes scapularis]	0.16	No hit
562	4-2_TriplEx25	1	ref|XP_001745774.1| hypothetical protein [Monosiga brevicollis MX1]	0.16	No hit
563	7-18_TriplEx25	1	ref|XP_001191172.1| PREDICTED: similar to endonuclease [S. purpuratus]	0.16	No hit
564	15-38_TriplEx25	1	ref|XP_002590756.1| hypothetical protein BRAFLDRAFT_78171 [B. floridae]	0.18	No hit
565	Contig35	11	gb|AAQ21039.1| ferritin [B. belcheri]	0.18	288000| ferritin
566	25-81_TriplEx25	1	ref|ZP_04608568.1| DNA polymerase III, alpha subunit [Micromonospora sp. ATCC 39149]	0.19	No hit
567	45-79_TriplEx25	1	ref|YP_003127412.1| small GTP-binding protein [Methanocaldococcus fervens AG86]	0.19	No hit
568	46-93_TriplEx25	1	ref|XP_002151908.1| cell polarity protein, putative [Penicillium marneffei ATCC 18224]	0.19	No hit
569	40-1_TriplEx25	1	emb|CAC41352.1| extracellular calcium sensing receptor precursor [Sparus aurata]	0.19	No hit
570	7-42_TriplEx25	1	ref|ZP_00208103.1| hypothetical protein [Magnetospirillum magnetotacticum]	0.21	No hit
571	1-7_TriplEx25	1	ref|XP_002607203.1| hypothetical protein BRAFLDRAFT_68000 [B. floridae]	0.21	No hit
572	8-4_TriplEx25	1	ref|XP_002604172.1| hypothetical protein BRAFLDRAFT_120397 [B. floridae]	0.21	No hit
573	15-60_TriplEx25	1	ref|XP_002371498.1| hypothetical protein TGME49_095590 [Toxoplasma gondii ME49]	0.21	No hit
574	13-89_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	0.21	No hit
575	17-36_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.22	No hit
576	8-23_TriplEx25	1	ref|XP_002603498.1| hypothetical protein BRAFLDRAFT_220050 [B. floridae]	0.23	No hit
577	35-29_TriplEx25	1	ref|XP_001718034.2| PREDICTED: hypothetical protein [Homo sapiens]	0.23	No hit
578	6-7_TriplEx25	1	ref|XP_001106300.1| PREDICTED: similar to TBC1 domain family [Macaca mulatta]	0.24	No hit
579	6-2_TriplEx25	1	ref|NP_064948.1| hypothetical protein AMV166 [Amsacta moorei entomopoxvirus]	0.24	No hit
580	47-23_TriplEx25	1	ref|ZP_03272867.1| conserved hypothetical protein [Arthrospira maxima CS-328]	0.25	No hit
581	23-55_TriplEx25	1	ref|ZP_01704707.1| transcriptional regulator, LysR family [Shewanella putrefaciens]	0.25	No hit
582	29-66_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.25	No hit
583	24-12_TriplEx25	1	emb|CBH17624.1| aminopeptidase, metallo-peptidase [Trypanosoma brucei gambiense]	0.25	No hit
584	Contig123	2	ref|XP_502130.1| YALI0C22297p [Yarrowia lipolytica]	0.26	No hit
585	2-42_TriplEx25	1	ref|XP_001866205.1| kek1 [Culex quinquefasciatus]	0.27	No hit
586	46-15_TriplEx25	1	ref|XP_001031432.1| Ubiquitin carboxyl-terminal hydrolase family [Tetrahymena thermophila]	0.27	No hit
587	Contig89	3	gb|EEH38949.1| predicted protein [Paracoccidioides brasiliensis Pb01]	0.27	No hit
588	13-44_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.27	No hit
589	Contig36	2	ref|ZP_05059876.1| glycosyl transferase [Verrucomicrobiae bacterium]	0.28	131007| hypothetical protein
590	9-14_TriplEx25	1	ref|XP_001349428.1| hypothetical protein [Plasmodium falciparum 3D7]	0.28	No hit
591	14-24_TriplEx25	1	ref|XP_001589805.1| predicted protein [Sclerotinia sclerotiorum 1980]	0.3	No hit
592	41-85_TriplEx25	1	ref|XP_002284001.1| PREDICTED: hypothetical protein [Vitis vinifera]	0.31	128859| similar to collagen, type XXII, alpha 1
593	41-94_TriplEx25	1	ref|XP_002122963.1| PREDICTED: similar to fer-1-like 3, myoferlin [C. intestinalis]	0.31	No hit
594	38-91_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	0.32	No hit
595	30-37_TriplEx25	1	ref|XP_001631903.1| predicted protein [Nematostella vectensis]	0.32	No hit
596	20-74_TriplEx25	1	gb|AAK58879.1|AF355375_1 putative reverse transcriptase [Takifugu rubripes]	0.32	No hit
597	38-2_TriplEx25	1	gb|AAI65718.1| Rars protein [Danio rerio]	0.32	No hit
598	27-34_TriplEx25	1	ref|XP_001788439.1| PREDICTED: similar to keratinocyte proline-rich protein [Bos taurus]	0.34	No hit
599	Contig12	2	ref|XP_628016.1| Cdc50p like membrane protein [Cryptosporidium parvum Iowa II]	0.36	No hit
600	2-73_TriplEx25	1	ref|XP_001250694.2| PREDICTED: similar to olfactory receptor Olr1307 [Bos taurus]	0.36	No hit

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
601	48-22_TriplEx25	1	ref|XP_002594058.1| hypothetical protein BRAFLDRAFT_118806 [B. floridae]	0.37	No hit
602	40-3_TriplEx25	1	ref|ZP_04643973.1| putative transcriptional antiterminator [Lactobacillus gasseri]	0.41	No hit
603	6-4_TriplEx25	1	ref|XP_001618589.1| hypothetical protein NEMVEDRAFT_v1g154106 [Nematostella vectensis]	0.41	No hit
604	41-55_TriplEx25	1	gb|ACX30979.1| heme lyase [Moneuplotes crassus]	0.41	No hit
605	45-8_TriplEx25	1	ref|YP_634887.1| hypothetical protein MXAN_6770 [Myxococcus xanthus DK 1622]	0.42	212802l similar to angiopoietin 2
606	36-35_TriplEx25	1	ref|YP_002250209.1| bacilysin biosynthesis oxidoreductase [Dictyoglomus thermophilum]	0.42	No hit
607	36-8_TriplEx25	1	ref|YP_001883839.1| hypothetical protein BH0412 [Borrelia hermsii DAH]	0.42	No hit
608	37-48_TriplEx25	1	ref|XP_002604527.1| hypothetical protein BRAFLDRAFT_79371 [B. floridae]	0.42	79371l hypothetical protein
609	17-31_TriplEx25	1	ref|XP_002589316.1| hypothetical protein BRAFLDRAFT_77769 [B. floridae]	0.42	97363l similar to DiGeorge syndrome critical region gene 8
610	22-18_TriplEx25	1	ref|XP_001524335.1| conserved hypothetical protein [Lodderomyces elongisporus]	0.42	No hit
611	47-56_TriplEx25	1	gb|EDL38452.1| mCG1050357 [Mus musculus]	0.42	No hit
612	32-6_TriplEx25	1	dbj|BAE01398.1| unnamed protein product [Macaca fascicularis]	0.42	No hit
613	22-35_TriplEx25	1	emb|CAA33503.1| dynein (515 AA) [Oncorhynchus mykiss]	0.45	No hit
614	Contig45	4	ref|ZP_01053327.1| conserved hypothetical protein [Polaribacter sp. MED152]	0.47	No hit
615	3-79_TriplEx25	1	ref|XP_727580.1| hypothetical protein [Plasmodium yoelii yoelii str. 17XNL]	0.47	No hit
616	10-87_TriplEx25	1	ref|XP_002610561.1| hypothetical protein BRAFLDRAFT_117847 [B. floridae]	0.47	117847l hypothetical protein
617	7-59_TriplEx25	1	ref|XP_002594138.1| hypothetical protein BRAFLDRAFT_211513 [B. floridae]	0.47	258920l glutamic-oxaloacetic transaminase 1
618	34-34_TriplEx25	1	ref|XP_002325252.1| autoinhibited calcium ATPase [Populus trichocarpa]	0.47	No hit
619	4-79_TriplEx25	1	ref|XP_001462519.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	0.47	No hit
620	Contig56	3	ref|NP_690429.1| Orf10 [Heliothis zea virus 1]	0.47	No hit
621	Contig130	2	gb|EEH53961.1| predicted protein [Micromonas pusilla CCMP1545]	0.47	No hit
622	Contig23	2	gb|EDZ70476.1| hypothetical protein AWRI1631_123600 [Saccharomyces cerevisiae]	0.47	No hit
623	18-72_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.49	No hit
624	38-71_TriplEx25	1	ref|ZP_03946832.1| glycosyltransferase involved in cell wall biogenesis [Atopobium vaginae]	0.54	No hit
625	40-42_TriplEx25	1	ref|ZP_03223188.1| putative integral membrane protein [Campylobacter jejuni]	0.54	No hit
626	38-93_TriplEx25	1	ref|NP_741585.1| hypothetical protein B0222.10 [Caenorhabditis elegans]	0.54	No hit
627	6-53_TriplEx25	1	dbj|BAH93963.1| Os07g0535300 [Oryza sativa Japonica Group]	0.54	No hit
628	30-66_TriplEx25	1	ref|ZP_06190965.1| probable microcin-H47 secretion [Serratia odorifera]	0.55	No hit
629	36-52_TriplEx25	1	ref|XP_591164.3| PREDICTED: similar to Interleukin-10 receptor[Bos taurus]	0.55	No hit
630	35-48_TriplEx25	1	ref|XP_002579334.1| hypothetical protein [Schistosoma mansoni]	0.55	No hit
631	45-40_TriplEx25	1	emb|CAL37898.1| hypothetical protein [synthetic construct]	0.55	No hit
632	28-12_TriplEx25	1	gb|EER05978.1| hypothetical protein Pmar_PMAR028165 [Perkinsus marinus]	0.58	No hit
633	12-48_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	0.6	No hit
634	16-16_TriplEx25	1	|XP_449067.1| unnamed protein product [Candida glabrata]	0.6	No hit
635	4-42_TriplEx25	1	gb|EEQ44094.1| allantoinase [Candida albicans WO-1]	0.62	No hit
636	18-52_TriplEx25	1	ref|XP_524874.2| PREDICTED: hornerin isoform 2 [Pan troglodytes]	0.63	No hit
637	6-30_TriplEx25	1	ref|XP_002313099.1| predicted protein [Populus trichocarpa]	0.7	No hit
638	6-91_TriplEx25	1	ref|XP_001636539.1| predicted protein [Nematostella vectensis]	0.7	No hit
639	41-16_TriplEx25	1	ref|XP_001350735.1| conserved Plasmodium protein [Plasmodium falciparum 3D7]	0.7	No hit
640	38-61_TriplEx25	1	gb|EER08109.1| hypothetical protein Pmar_PMAR014873 [Perkinsus marinus]	0.71	No hit
641	47-90_TriplEx25	1	ref|XP_001745460.1| hypothetical protein [Monosiga brevicollis MX1]	0.72	No hit
642	35-22_TriplEx25	1	ref|ZP_04873591.1| Peptidase M16 inactive domain family [Aciduliprofundum boonei]	0.72	No hit
643	35-44_TriplEx25	1	ref|YP_740790.1| Ymf77 [Tetrahymena paravorax]	0.72	No hit
644	37-64_TriplEx25	1	ref|YP_003304512.1| TonB-dependent receptor [Sulfurospirillum deleyianum DSM 6946]	0.72	No hit
645	23-61_TriplEx25	1	ref|YP_001545261.1| hypothetical protein Haur_2495 [Herpetosiphon aurantiacus ATCC]	0.72	No hit
646	22-34_TriplEx25	1	ref|XP_963919.1| hypothetical protein NCU07468 [Neurospora crassa OR74A]	0.72	No hit
647	36-3_TriplEx25	1	ref|XP_797849.1| PREDICTED: similar to conserved hypothetical protein [S. purpuratus]	0.72	No hit
648	46-68_TriplEx25	1	ref|XP_742315.1| hypothetical protein [Plasmodium chabaudi chabaudi]	0.72	No hit
649	36-31_TriplEx25	1	ref|XP_715091.1| hypothetical YFW family protein 8 [Candida albicans SC5314]	0.72	No hit
650	23-83_TriplEx25	1	ref|XP_002608992.1| hypothetical protein BRAFLDRAFT_84798 [B. floridae]	0.72	No hit
651	35-10_TriplEx25	1	ref|XP_002448990.1| hypothetical protein SORBIDRAFT_05g002970 [Sorghum bicolor]	0.72	No hit
652	31-58_TriplEx25	1	ref|XP_001924801.1| PREDICTED: similar to arylacetamide deacetylase-like 1 [Sus scrofa]	0.72	No hit
653	45-42_TriplEx25	1	ref|XP_001322646.1| hypothetical protein [Trichomonas vaginalis G3]	0.72	No hit
654	30-9_TriplEx25	1	gb|EEU44269.1| hypothetical protein NECHADRAFT_96107 [Nectria haematococca]	0.72	No hit
655	20-88_TriplEx25	1	gb|ABM68196.1| ZNF483 [Lagothrix lagotricha]	0.72	No hit
656	19-55_TriplEx25	1	ref|XP_001717389.1| PREDICTED: hypothetical protein [Homo sapiens]	0.73	No hit
657	Contig127	2	ref|ZP_06190130.1| TAP domain-containing protein [Serratia odorifera 4Rx13]	0.74	No hit
658	28-30_TriplEx25	1	ref|ZP_05759429.1| hypothetical protein BacD2_14193 [Bacteroides sp. D2]	0.76	No hit
659	4-27_TriplEx25	1	ref|ZP_04755212.1| hydrogenase maturation protein HypF [Francisella philomiragia]	0.81	No hit
660	19-87_TriplEx25	1	ref|YP_001154523.1| RNA polymerase, insert [Pyrobaculum arsenaticum DSM 13514]	0.83	No hit
661	49-82_TriplEx25	1	ref|XP_738216.1| hypothetical protein [Plasmodium chabaudi chabaudi]	0.83	No hit
662	18-6_TriplEx25	1	emb|CAM73970.1| hypothetical protein [Magnetospirillum gryphiswaldense MSR-1]	0.83	No hit
663	28-90_TriplEx25	1	ref|XP_002272204.1| PREDICTED: hypothetical protein [Vitis vinifera]	0.84	No hit
664	14-16_TriplEx25	1	gb|EDL96768.1| rCG50930 [Rattus norvegicus]	0.88	No hit
665	10-50_TriplEx25	1	ref|XP_664986.1| hypothetical protein [Cryptosporidium hominis TU502]	0.9	No hit
666	42-79_TriplEx25	1	ref|YP_001771469.1| 5-oxoprolinase (ATP-hydrolyzing) [Methylobacterium sp. 4-46]	0.91	No hit
667	6-1_TriplEx25	1	ref|XP_002649764.1| carboxy-terminal domain (CTD) phosphatase [Enterocytozoon bieneusi]	0.91	No hit
668	6-40_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	0.91	No hit
669	5-53_TriplEx25	1	ref|XP_002546536.1| conserved hypothetical protein [Candida tropicalis MYA-3404]	0.91	No hit
670	38-22_TriplEx25	1	gb|ACX30944.1| heme lyase [Euplotes minuta]	0.92	No hit
671	38-9_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.92	No hit
672	38-45_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	0.92	No hit
673	40-2_TriplEx25	1	gb|ABU89524.1| maturase K [Cestrum parqui]	0.92	No hit
674	29-48_TriplEx25	1	ref|XP_001440474.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	0.93	No hit
675	29-44_TriplEx25	1	gb|AAW76840.1| ISxac3 transposase [Xanthomonas oryzae pv. oryzae KACC10331]	0.93	No hit
676	45-80_TriplEx25	1	ref|XP_972454.1| PREDICTED: similar to AGAP001879-PA [Tribolium castaneum]	0.94	No hit
677	22-14_TriplEx25	1	ref|XP_743103.1| hypothetical protein [Plasmodium chabaudi chabaudi]	0.94	No hit
678	46-12_TriplEx25	1	ref|XP_002597387.1| hypothetical protein BRAFLDRAFT_69334 [B. floridae]	0.94	No hit
679	21-55_TriplEx25	1	ref|XP_002342505.1| PREDICTED: hypothetical protein XP_002342505 [Homo sapiens]	0.94	No hit
680	22-42_TriplEx25	1	ref|XP_001821663.1| hypothetical protein [Aspergillus oryzae RIB40]	0.94	No hit
681	47-52_TriplEx25	1	ref|XP_001731949.1| hypothetical protein MGL_1217 [Malassezia globosa CBS 7966]	0.94	No hit
682	20-78_TriplEx25	1	gb|EEC70577.1| hypothetical protein OsI_01776 [Oryza sativa Indica Group]	0.94	No hit
683	26-7_TriplEx25	1	ref|XP_001920228.1| similar to parathyroid hormone receptor 1 [Danio rerio]	0.99	No hit
684	1-59_TriplEx25	1	ref|YP_001687042.1| hypothetical protein SGHV094 [Glossina pallidipes]	1	No hit
685	8-12_TriplEx25	1	ref|YP_001661540.1| membrane protein [Streptomyces sp. HK1]	1	No hit
686	15-70_TriplEx25	1	ref|XP_001984648.1| GH14913 [Drosophila grimshawi]	1	No hit
687	15-62_TriplEx25	1	ref|XP_001350472.1| conserved Plasmodium protein [Plasmodium falciparum 3D7]	1	No hit
688	11-13_TriplEx25	1	gb|ABF97426.1| transposon protein, putative, unclassified [Oryza sativa]	1	No hit
689	3-67_TriplEx25	1	ref|XP_002590756.1| hypothetical protein BRAFLDRAFT_78171 [B. floridae]	1.1	No hit
690	19-34_TriplEx25	1	ref|XP_001907254.1| unnamed protein product [Podospora anserina]	1.1	No hit
691	48-64_TriplEx25	1	ref|XP_001714476.1| PREDICTED: hypothetical protein [Homo sapiens]	1.1	No hit
692	17-7_TriplEx25	1	ref|XP_001615766.1| hypothetical protein [Plasmodium vivax Sal-1]	1.1	No hit
693	49-43_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	1.1	No hit
694	3-53_TriplEx25	1	dbj|BAF57253.1| NADH dehydrogenase subunit F [Carex albata]	1.1	No hit
695	6-62_TriplEx25	1	ref|ZP_04740464.1| putative type I site-specific deoxyribonuclease [Neisseria gonorrhoeae]	1.2	No hit
696	46-91_TriplEx25	1	ref|ZP_03271633.1| sulfotransferase [Arthrospira maxima CS-328]	1.2	No hit
697	38-41_TriplEx25	1	ref|ZP_00545232.1| hypothetical protein EchaDRAFT_0032 [Ehrlichia chaffeensis]	1.2	No hit
698	15-79_TriplEx25	1	ref|YP_003009201.1| binding-protein-dependent transport component [Paenibacillus sp.]	1.2	No hit
699	29-4_TriplEx25	1	ref|XP_431583.1| hypothetical protein CaO19.1472 [Candida albicans SC5314]	1.2	No hit
700	25-92_TriplEx25	1	ref|XP_002638375.1| Hypothetical protein CBG18580 [Caenorhabditis briggsae]	1.2	No hit

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
701	6-5_TriplEx25	1	ref|XP_002605014.1| hypothetical protein BRAFLDRAFT_124129 [B. floridae]	1.2	111906| gastric intrinsic factor
702	32-74_TriplEx25	1	ref|XP_002526779.1| leucine-rich repeat containing protein, putative [Ricinus communis]	1.2	No hit
703	45-12_TriplEx25	1	ref|XP_001980288.1| GG19589 [Drosophila erecta]	1.2	No hit
704	47-27_TriplEx25	1	ref|XP_001952425.1| PREDICTED: similar to CG9601 CG9601-PA [Acyrthosiphon pisum]	1.2	No hit
705	36-58_TriplEx25	1	ref|XP_001316052.1| hypothetical protein [Trichomonas vaginalis G3]	1.2	No hit
706	32-53_TriplEx25	1	ref|NP_245873.1| gamma-glutamyl phosphate reductase [Pasteurella multocida]	1.2	No hit
707	25-87_TriplEx25	1	gb|EDL35754.1| mCG145705 [Mus musculus]	1.2	No hit
708	20-10_TriplEx25	1	gb|EAZ16390.1| hypothetical protein OsJ_31855 [Oryza sativa Japonica]	1.2	No hit
709	47-25_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	1.2	No hit
710	44-65_TriplEx25	1	gb|AAI69288.2| solute carrier family 12, member 2 [synthetic construct]	1.2	No hit
711	47-88_TriplEx25	1	emb|CBH09980.1| hypothetical protein, unlikely [Trypanosoma brucei]	1.2	No hit
712	37-47_TriplEx25	1	ref|YP_003073655.1| modular polyketide synthase, type I PKS [Teredinibacter turnerae]	1.2	No hit
713	12-74_TriplEx25	1	ref|ZP_03626563.1| autotransporter-associated beta strand repeat protein [bacterium Ellin514]	1.3	No hit
714	28-19_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	1.3	No hit
715	1-50_TriplEx25	1	ref|XP_001613253.1| hypothetical protein [Plasmodium vivax Sal-1]	1.3	No hit
716	12-79_TriplEx25	1	ref|NP_001132397.1| hypothetical protein LOC100193843 [Zea mays]	1.3	No hit
717	Contig134	2	ref|NP_001089881.1| hypothetical protein LOC734948 [Xenopus laevis]	1.3	No hit
718	18-46_TriplEx25	1	ref|XP_001684662.1| CDC16 [Leishmania major strain Friedlin]	1.4	No hit
719	17-90_TriplEx25	1	ref|XP_001456265.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	1.4	No hit
720	19-8_TriplEx25	1	ref|XP_001439622.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	1.4	No hit
721	48-18_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	1.4	89299| hypothetical protein
722	Contig11	3	gb|AAB17561.1| LRG5 [Chlamydomonas reinhardtii	1.4	No hit
723	30-11_TriplEx25	1	ref|ZP_06077970.1| conserved hypothetical protein [Bacteroides sp. 2_1_33B]	1.6	No hit
724	47-94_TriplEx25	1	ref|ZP_04150928.1| YcnI (Conserved membrane protein) [Bacillus pseudomycoides DSM]	1.6	No hit
725	37-81_TriplEx25	1	ref|YP_802916.1| Pgi [Buchnera aphidicola str. Cc (Cinara cedri)]	1.6	No hit
726	31-44_TriplEx25	1	ref|YP_002572082.1| glycosyl transferase family 2 [Anaerocellum thermophilum]	1.6	No hit
727	36-59_TriplEx25	1	ref|YP_001738254.1| hypothetical protein TRQ2_0211 [Thermotoga sp. RQ2]	1.6	No hit
728	30-14_TriplEx25	1	ref|YP_001313280.1| polar amino acid ABC transporter [Sinorhizobium medicae WSM419]	1.6	No hit
729	42-38_TriplEx25	1	ref|XP_868375.1| PREDICTED: similar to active BCR-related gene [Canis familiaris]	1.6	No hit
730	40-27_TriplEx25	1	ref|XP_002612944.1| hypothetical protein BRAFLDRAFT_213400 [B. floridae]	1.6	No hit
731	25-67_TriplEx25	1	ref|XP_002609806.1| hypothetical protein BRAFLDRAFT_219541 [B. floridae]	1.6	No hit
732	31-81_TriplEx25	1	ref|XP_002603466.1| hypothetical protein BRAFLDRAFT_153964 [B. floridae]	1.6	No hit
733	22-74_TriplEx25	1	ref|XP_002424907.1| galactokinase, putative [Pediculus humanus corporis]	1.6	No hit
734	41-63_TriplEx25	1	ref|XP_001849742.1| conserved hypothetical protein [Culex quinquefasciatus]	1.6	No hit
735	45-68_TriplEx25	1	ref|NP_177694.1| leucine-rich repeat family protein [Arabidopsis thaliana]	1.6	No hit
736	41-51_TriplEx25	1	gb|ABV69452.1| RNA polymerase II largest subunit [Endocarpon adscendens]	1.6	No hit
737	29-20_TriplEx25	1	gb|ABV68938.1| nonstructural polyprotein [Getah virus]	1.6	No hit
738	36-76_TriplEx25	1	gb|AAP35717.1| unknown [Pseudomonas aeruginosa]	1.6	No hit
739	38-92_TriplEx25	1	emb|CBI19832.1| unnamed protein product [Vitis vinifera]	1.6	No hit
740	46-10_TriplEx25	1	emb|CAL53302.1| anaphase promoting complex subunit 3 [Ostreococcus tauri]	1.6	No hit
741	12-17_TriplEx25	1	ref|ZP_01619372.1| mannosyltransferase [Lyngbya sp. PCC 8106]	1.7	No hit
742	28-80_TriplEx25	1	ref|YP_866204.1| putative PAS/PAC sensor protein [Magnetococcus sp. MC-1]	1.7	No hit
743	Contig159	2	ref|YP_823639.1| acriflavin resistance protein [Solibacter usitatus Ellin6076]	1.7	No hit
744	10-53_TriplEx25	1	sp|Q5ZKN5.2|FA53A_CHICK RecName: Full=Protein FAM53A	1.8	No hit
745	34-81_TriplEx25	1	ref|YP_002787913.1| putative nitrilase [Deinococcus deserti VCD115]	1.8	No hit
746	33-57_TriplEx25	1	ref|XP_002589136.1| hypothetical protein BRAFLDRAFT_213755 [B. floridae]	1.8	No hit
747	9-11_TriplEx25	1	ref|XP_002558489.1| Pc13g00400 [Penicillium chrysogenum Wisconsin 54-1255]	1.8	No hit
748	Contig37	3	ref|XP_001919369.1| similar to asparagine-linked glycosylation 10 homolog [Danio rerio]	1.8	No hit
749	49-89_TriplEx25	1	ref|XP_001746548.1| hypothetical protein [Monosiga brevicollis MX1]	1.8	No hit
750	Contig48	2	ref|NP_989302.2| solute carrier family 34 [Xenopus tropicalis]	1.8	No hit
751	10-13_TriplEx25	1	gb|EEH17837.1| high affinity nitrate transporter NrtB [Paracoccidioides brasiliensis]	1.8	131007| hypothetical protein
752	Contig42	3	gb|EEE31760.1| phosphatidylinositol 3-, 4-kinase domain-containing protein [Toxoplasma gondii]	1.8	No hit
753	2-49_TriplEx25	1	gb|EEC10032.1| CEBPA: CCAAT/enhancer-binding protein alpha [Ixodes scapularis]	1.8	No hit
754	48-74_TriplEx25	1	gb|ACI67429.1| Steroid receptor RNA activator 1 [Salmo salar]	1.8	No hit
755	2-8_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	1.8	No hit
756	8-73_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	1.8	No hit
757	15-24_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	1.9	No hit
758	42-81_TriplEx25	1	ref|YP_002382590.1| Acid shock protein precursor [Escherichia fergusonii]	2	No hit
759	42-41_TriplEx25	1	ref|YP_001955952.1| CoA enzyme activase [uncultured Termite group 1 bacterium]	2	No hit
760	6-80_TriplEx25	1	ref|XP_817857.1| mucin TcMUCII [Trypanosoma cruzi strain CL Brener]	2	No hit
761	41-29_TriplEx25	1	ref|XP_503926.1| YALI0E14036p [Yarrowia lipolytica]	2	No hit
762	41-17_TriplEx25	1	ref|XP_001009220.1| hypothetical protein TTHERM_00554300 [Tetrahymena thermophila]	2	No hit
763	35-12_TriplEx25	1	ref|ZP_05631212.1| exonuclease SBCC [Fusobacterium gonidiaformans ATCC 25563]	2.1	No hit
764	38-72_TriplEx25	1	ref|ZP_05068787.1| conserved hypothetical protein [Candidatus Pelagibacter sp. HTCC7211]	2.1	No hit
765	45-89_TriplEx25	1	ref|ZP_04591040.1| hypothetical protein Psyrpo1_27451 [Pseudomonas syringae]	2.1	No hit
766	46-29_TriplEx25	1	ref|ZP_04080414.1| Membrane protein, possible ABC transporter [Bacillus thuringiensis]	2.1	No hit
767	38-66_TriplEx25	1	ref|ZP_01812001.1| hypothetical protein VSWAT3_25889 [Vibrionales bacterium SWAT-3]	2.1	No hit
768	36-81_TriplEx25	1	ref|YP_001548975.1| hypothetical protein MmarC6_0928 [Methanococcus maripaludis C6]	2.1	No hit
769	20-83_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	2.1	No hit
770	25-28_TriplEx25	1	ref|XP_002400981.1| suppressor protein SRP40, putative [Ixodes scapularis]	2.1	No hit
771	29-38_TriplEx25	1	ref|XP_002313448.1| predicted protein [Populus trichocarpa]	2.1	No hit
772	25-60_TriplEx25	1	ref|XP_001922408.1| PREDICTED: similar to Zinc finger protein 84 [Danio rerio]	2.1	No hit
773	9-61_TriplEx25	1	ref|XP_001625994.1| predicted protein [Nematostella vectensis]	2.1	No hit
774	20-15_TriplEx25	1	ref|NP_781592.1| benzylsuccinate synthase activating enzyme [Clostridium tetani]	2.1	No hit
775	30-73_TriplEx25	1	ref|NP_001027100.1| CG33791, isoform B [Drosophila melanogaster]	2.1	No hit
776	35-39_TriplEx25	1	gb|EET02650.1| Hypothetical protein GL50581_37 [Giardia intestinalis ATCC 50581]	2.1	No hit
777	29-71_TriplEx25	1	gb|EER38546.1| conserved hypothetical protein [Ajellomyces capsulatus H143]	2.1	No hit
778	35-21_TriplEx25	1	gb|EDL38261.1| EGF-like module containing, mucin-like, hormone receptor-like [Mus musculus]	2.1	No hit
779	29-19_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	2.1	No hit
780	47-20_TriplEx25	1	emb|CBH14329.1| hypothetical protein, unlikely [Trypanosoma brucei gambiense]	2.1	No hit
781	36-72_TriplEx25	1	emb|CAX14567.1| novel protein [Danio rerio]	2.1	No hit
782	Contig136	4	ref|YP_002720837.1| hypothetical protein BHWA1_00637 [Brachyspira hyodysenteriae]	2.2	No hit
783	41-45_TriplEx25	1	ref|XP_672303.1| Pb-fam-2 protein [Plasmodium berghei strain ANKA]	2.2	No hit
784	27-6_TriplEx25	1	ref|XP_002609842.1| hypothetical protein BRAFLDRAFT_122138 [B. floridae]	2.2	No hit
785	10-4_TriplEx25	1	ref|NP_616851.1| hypothetical protein MA1927 [Methanosarcina acetivorans C2A]	2.2	No hit
786	9-8_TriplEx25	1	ref|YP_266861.1| hypothetical protein CPS_0093 [Colwellia psychrerythraea 34H]	2.3	No hit
787	7-76_TriplEx25	1	ref|YP_161509.1| hypothetical protein BGP223 [Borrelia garinii PBi]	2.3	No hit
788	4-47_TriplEx25	1	ref|YP_002311857.1| FAD dependent oxidoreductase [Shewanella piezotolerans WP3]	2.3	No hit
789	10-81_TriplEx25	1	ref|XP_730161.1| hypothetical protein [Plasmodium yoelii yoelii str. 17XNL]	2.3	No hit
790	8-42_TriplEx25	1	ref|XP_002568325.1| Pc21g12990 [Penicillium chrysogenum Wisconsin 54-1255]	2.3	No hit
791	10-18_TriplEx25	1	ref|XP_002431845.1| oxysterol-binding protein 3, putative [Pediculus humanus corporis]	2.3	No hit
792	7-65_TriplEx25	1	ref|XP_002307025.1| cc-nbs-lrr resistance protein [Populus trichocarpa]	2.3	No hit
793	34-27_TriplEx25	1	ref|XP_002131762.1| PREDICTED: hypothetical protein isoform 1 [C. intestinalis]	2.3	No hit
794	3-16_TriplEx25	1	ref|XP_001717872.1| PREDICTED: hypothetical protein [Homo sapiens]	2.3	No hit
795	Contig27	2	ref|XP_001200401.1| PREDICTED: similar to CRH receptor 2 [S. purpuratus]	2.3	No hit
796	5-14_TriplEx25	1	ref|XP_001008467.1| hypothetical protein TTHERM_00022820 [Tetrahymena thermophila]	2.3	No hit
797	7-96_TriplEx25	1	gb|EER39117.1| conserved hypothetical protein [Ajellomyces capsulatus H143]	2.3	No hit
798	10-52_TriplEx25	1	emb|CAG13138.1| unnamed protein product [Tetraodon nigroviridis]	2.3	No hit
799	49-37_TriplEx25	1	ref|YP_002432678.1| transport system permease protein [Desulfatibacillum alkenivorans]	2.4	No hit
800	17-84_TriplEx25	1	ref|XP_651504.2| hypothetical protein [Entamoeba histolytica HM-1:IMSS]	2.4	No hit

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
801	6-3_TriplEx25	1	ref|ZP_03799410.1| hypothetical protein COPCOM_01667 [Coprococcus comes ATCC 27758]	2.7	79372l hypothetical protein
802	40-54_TriplEx25	1	ref|ZP_03777793.1| hypothetical protein CLOHYLEM_04847 [Clostridium hylemonae]	2.7	No hit
803	42-12_TriplEx25	1	ref|ZP_03054584.1| polyketide synthase subunit [Bacillus pumilus ATCC 7061]	2.7	No hit
804	35-88_TriplEx25	1	ref|ZP_02634610.1| M protein trans-acting positive regulator (MGA) [Clostridium perfringens]	2.7	No hit
805	30-32_TriplEx25	1	ref|YP_822533.1| inositol-3-phosphate synthase [Solibacter usitatus Ellin6076]	2.7	No hit
806	36-78_TriplEx25	1	ref|YP_002721824.1| hypothetical protein BHWA1_01650 [Brachyspira hyodysenteriae]	2.7	No hit
807	20-41_TriplEx25	1	ref|XP_724053.1| hypothetical protein [Plasmodium yoelii yoelii str. 17XNL]	2.7	No hit
808	25-34_TriplEx25	1	ref|XP_723781.1| hypothetical protein [Plasmodium yoelii yoelii str. 17XNL]	2.7	No hit
809	38-58_TriplEx25	1	ref|XP_002589136.1| hypothetical protein BRAFLDRAFT_213755 [B. floridae]	2.7	No hit
810	45-69_TriplEx25	1	ref|XP_002479017.1| glycosyl hydrolase, putative [Talaromyces stipitatus ATCC 10500]	2.7	No hit
811	35-54_TriplEx25	1	ref|XP_002417245.1| leucine-rich lrepeat IFA/LPF family protein [Candida dubliniensis CD36]	2.7	No hit
812	5-40_TriplEx25	1	ref|XP_002344586.1| PREDICTED: hypothetical protein, partial [Homo sapiens]	2.7	No hit
813	23-8_TriplEx25	1	ref|XP_002061356.1| GK20876 [Drosophila willistoni]	2.7	No hit
814	30-48_TriplEx25	1	ref|XP_001987220.1| GH21800 [Drosophila grimshawi]	2.7	No hit
815	23-46_TriplEx25	1	ref|XP_001009658.1| hypothetical protein TTHERM_00155300 [Tetrahymena thermophila]	2.7	No hit
816	29-25_TriplEx25	1	ref|NP_963557.1| hypothetical protein NEQ266 [Nanoarchaeum equitans Kin4-M]	2.7	No hit
817	35-30_TriplEx25	1	gb|EEQ27913.1| predicted protein [Microsporum canis CBS 113480]	2.7	No hit
818	20-9_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	2.7	No hit
819	22-77_TriplEx25	1	gb|AAM47018.1| ionotropic GABA receptor beta subunit 1b [Homarus americanus]	2.7	No hit
820	36-49_TriplEx25	1	gb|AAB19113.1| longevity assurance factor [Schizosaccharomyces pombe]	2.7	No hit
821	47-72_TriplEx25	1	emb|CBH10617.1| T. brucei spp.-specific protein [Trypanosoma brucei gambiense]	2.7	No hit
822	22-38_TriplEx25	1	emb|CAE18134.1| replication-associated protein [Tomato yellow leaf curl China virus]	2.7	No hit
823	Contig161	2	ref|XP_001962187.1| GF14565 [Drosophila ananassae]	2.8	No hit
824	27-21_TriplEx25	1	ref|XP_397205.2| PREDICTED: similar to suppressor of Hairy wing [Apis mellifera]	2.9	No hit
825	27-91_TriplEx25	1	ref|XP_002131995.1| PREDICTED: similar to arsenate resistance protein 2 [C. intestinalis]	2.9	No hit
826	27-93_TriplEx25	1	ref|XP_001359988.2| GA20291 [Drosophila pseudoobscura pseudoobscura]	2.9	No hit
827	27-23_TriplEx25	1	ref|NP_896529.1| hypothetical protein SYNW0434 [Synechococcus sp. WH 8102]	2.9	No hit
828	26-15_TriplEx25	1	gb|AAZ43915.2| putative ABC transporter, ATP-binding protein [Mycoplasma synoviae]	2.9	No hit
829	Contig51	2	ref|YP_073295.1| NADH dehydrogenase subunit 1 [Pachypsylla venusta]	3	No hit
830	9-40_TriplEx25	1	ref|XP_861680.1| PREDICTED: similar to cytoplasmic beta-actin isoform 4 [Canis familiaris]	3	No hit
831	1-13_TriplEx25	1	ref|XP_381322.1| hypothetical protein FG01146.1 [Gibberella zeae PH-1]	3	No hit
832	Contig52	2	ref|XP_002632079.1| Hypothetical protein CBG17045 [Caenorhabditis briggsae]	3	No hit
833	10-61_TriplEx25	1	ref|XP_002629834.1| Hypothetical protein CBG18722 [Caenorhabditis briggsae]	3	No hit
834	34-15_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	3	No hit
835	33-50_TriplEx25	1	ref|XP_002581664.1| hypothetical protein [Schistosoma mansoni]	3	No hit
836	10-49_TriplEx25	1	ref|XP_002347432.1| PREDICTED: similar to hCG1820441 [Homo sapiens]	3	No hit
837	13-58_TriplEx25	1	ref|XP_002347432.1| PREDICTED: similar to hCG1820441 [Homo sapiens]	3	No hit
838	34-20_TriplEx25	1	ref|XP_002166144.1| PREDICTED: similar to predicted protein [Hydra magnipapillata]	3	No hit
839	Contig54	2	ref|XP_001714788.2| PREDICTED: hypothetical protein [Homo sapiens]	3	No hit
840	12-24_TriplEx25	1	ref|XP_001314546.1| hypothetical protein [Trichomonas vaginalis G3]	3	No hit
841	34-78_TriplEx25	1	ref|XP_001012069.1| hypothetical protein TTHERM_00985180 [Tetrahymena thermophila]	3	No hit
842	9-12_TriplEx25	1	dbj|BAI50623.1| caseinolytic protease C [Plasmodium simiovale]	3	25892l glutamic-oxaloacetic transaminase 1
843	18-1_TriplEx25	1	ref|ZP_03938011.1| glycoside hydrolase family 25 [Lactobacillus brevis subsp. Gravesensis]	3.1	No hit
844	19-10_TriplEx25	1	ref|ZP_01695018.1| hypothetical protein M23134_01304 [Microscilla marina ATCC 23134]	3.1	No hit
845	49-42_TriplEx25	1	ref|XP_447533.1| hypothetical protein CAGL0I06512g [Candida glabrata CBS138]	3.1	No hit
846	49-29_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	3.1	No hit
847	3-19_TriplEx25	1	ref|XP_002336964.1| predicted protein [Populus trichocarpa]	3.1	No hit
848	48-54_TriplEx25	1	ref|XP_001405735.1| predicted protein [Magnaporthe grisea 70-15]	3.1	No hit
849	48-79_TriplEx25	1	ref|XP_001017090.2| Plasmid Maintenance Protein [Tetrahymena thermophila]	3.1	No hit
850	49-25_TriplEx25	1	gb|EER44030.1| leucine rich repeat domain-containing protein [Ajellomyces capsulatus]	3.1	No hit
851	4-89_TriplEx25	1	emb|CAN71287.1| hypothetical protein [Vitis vinifera]	3.1	No hit
852	15-5_TriplEx25	1	ref|ZP_06127198.1| tagatose 6-phosphate kinase [Providencia rettgeri DSM 1131]	3.3	No hit
853	38-47_TriplEx25	1	ref|ZP_05783654.1| trap transporter, dctm subunit [Citreicella sp. SE45]	3.5	No hit
854	6-20_TriplEx25	1	ref|ZP_04582040.1| conserved hypothetical protein [Helicobacter bilis ATCC 43879]	3.5	No hit
855	42-80_TriplEx25	1	ref|ZP_02370937.1| YD repeat protein [Burkholderia thailandensis TXDOH]	3.5	No hit
856	42-54_TriplEx25	1	ref|ZP_02166549.1| valyl-tRNA synthetase [Hoeflea phototrophica DFL-43]	3.5	No hit
857	40-33_TriplEx25	1	ref|YP_002413093.1| Mannosyl transferase WbaC [Escherichia coli UMN026]	3.5	No hit
858	41-80_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	3.5	No hit
859	42-15_TriplEx25	1	ref|XP_001942566.1| PREDICTED: similar to 2,5-phosphodiesterase [Acyrthosiphon pisum]	3.5	No hit
860	28-55_TriplEx25	1	ref|XP_001747731.1| hypothetical protein [Monosiga brevicollis MX1]	3.5	No hit
861	42-11_TriplEx25	1	ref|XP_001504889.2| PREDICTED: similar to olfactory receptor [Equus caballus]	3.5	No hit
862	42-27_TriplEx25	1	ref|XP_001220065.1| hypothetical protein CHGG_00844 [Chaetomium globosum CBS 148.51]	3.5	No hit
863	44-87_TriplEx25	1	ref|NP_519257.1| hypothetical protein RSc1136 [Ralstonia solanacearum GMI1000]	3.5	No hit
864	40-90_TriplEx25	1	emb|CBI27244.1| unnamed protein product [Vitis vinifera]	3.5	No hit
865	32-25_TriplEx25	1	ref|ZP_05711740.1| conserved hypothetical protein [Listeria monocytogenes FSL R2-503]	3.6	No hit
866	46-28_TriplEx25	1	ref|ZP_05005080.1| hypothetical protein SSCG_02407 [Streptomyces clavuligerus]	3.6	No hit
867	30-43_TriplEx25	1	ref|ZP_04394435.1| conserved hypothetical protein [Geobacillus sp. Y412MC52]	3.6	No hit
868	22-46_TriplEx25	1	ref|ZP_01061350.1| type IV site-specific deoxyribonuclease [Leeuwenhoekiella blandensis]	3.6	88967l similar to kelch-like 12 isoform 1
869	45-48_TriplEx25	1	ref|ZP_00056111.1| COG0196: FAD synthase [Magnetospirillum magnetotacticum MS-1]	3.6	No hit
870	35-90_TriplEx25	1	ref|YP_003069050.1| dihydrolipoamide dehydrogenase [Methylobacterium extorquens DM4]	3.6	No hit
871	46-20_TriplEx25	1	ref|XP_765158.1| hypothetical protein [Theileria parva strain Muguga]	3.6	No hit
872	37-11_TriplEx25	1	ref|XP_002591315.1| hypothetical protein BRAFLDRAFT_76766 [B. floridae]	3.6	No hit
873	25-39_TriplEx25	1	ref|XP_002589795.1| hypothetical protein BRAFLDRAFT_125897 [B. floridae]	3.6	No hit
874	35-3_TriplEx25	1	ref|XP_002588568.1| hypothetical protein BRAFLDRAFT_110741 [B. floridae]	3.6	No hit
875	47-7_TriplEx25	1	ref|XP_002417500.1| pab1p-dependent poly(a)-nuclease [Candida dubliniensis CD36]	3.6	No hit
876	35-36_TriplEx25	1	ref|XP_002342485.1| PREDICTED: hypothetical protein XP_002342485 [Homo sapiens]	3.6	No hit
877	20-22_TriplEx25	1	ref|XP_002263903.1| PREDICTED: hypothetical protein [Vitis vinifera]	3.6	No hit
878	25-93_TriplEx25	1	ref|XP_002087466.1| GE17067 [Drosophila yakuba]	3.6	No hit
879	36-87_TriplEx25	1	ref|XP_001686840.1| protein kinase [Leishmania major strain Friedlin]	3.6	12484l slit (Drosophila) homolog 2
880	22-76_TriplEx25	1	ref|XP_001663511.1| receptor protein tyrosine phosphatase, putative [Aedes aegypti]	3.6	No hit
881	23-68_TriplEx25	1	ref|XP_001608552.1| variable surface protein Vir27 [Plasmodium vivax Sal-1]	3.6	No hit
882	32-45_TriplEx25	1	ref|XP_001454859.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	3.6	No hit
883	30-61_TriplEx25	1	ref|XP_001439926.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	3.6	No hit
884	20-26_TriplEx25	1	ref|XP_001162008.1| PREDICTED: hypothetical protein [Pan troglodytes]	3.6	No hit
885	35-72_TriplEx25	1	gb|EEC68377.1| hypothetical protein Os|_36516 [Oryza sativa Indica Group]	3.6	No hit
886	46-27_TriplEx25	1	gb|ABB88698.1| P-type ATPase [Dunaliella salina]	3.6	No hit
887	35-56_TriplEx25	1	gb|AAP41672.1| maturase K [Allocasuarina decaisneana	3.6	No hit
888	45-81_TriplEx25	1	gb|AAO12119.1| aminoglycoside adenylyltransferase [Streptococcus oralis]	3.6	No hit
889	47-16_TriplEx25	1	gb|AAF05916.1| delta-12 oleic acid desaturase-like protein [Momordica charantia]	3.6	No hit
890	31-39_TriplEx25	1	emb|CBI34341.1| unnamed protein product [Vitis vinifera]	3.6	No hit
891	20-11_TriplEx25	1	emb|CAJ73019.1| similar to histidine kinase [Candidatus Kuenenia stuttgartiensis]	3.6	No hit
892	45-55_TriplEx25	1	dbj|BAG55488.1| receptor-type protein tyrosine kinase [Monosiga ovata]	3.6	No hit
893	38-96_TriplEx25	1	ref|NP_492372.1| Helix Loop Helix family member (hlh-16) [Caenorhabditis elegans]	3.7	No hit
894	28-40_TriplEx25	1	ref|XP_503168.1| YALI0D22891p [Yarrowia lipolytica]	3.8	No hit
895	27-72_TriplEx25	1	ref|XP_001733702.1| hypothetical protein [Entamoeba dispar SAW760]	3.8	No hit
896	28-84_TriplEx25	1	gb|ACA53494.1| olfactory receptor Olr135 (predicted) [Callicebus moloch]	3.8	No hit
897	16-43_TriplEx25	1	ref|XP_001875110.1| predicted protein [Laccaria bicolor S238N-H82]	3.9	No hit
898	12-41_TriplEx25	1	ref|XP_001586017.1| hypothetical protein SS1G_13110 [Sclerotinia sclerotiorum 1980]	3.9	13334l similar to macrophage mennose receptor 1 precursor
899	12-87_TriplEx25	1	ref|XP_001014246.1| hypothetical protein TTHERM_00227150 [Tetrahymena thermophila]	3.9	No hit
900	2-4_TriplEx25	1	gb|EEU05501.1| Gas2p [Saccharomyces cerevisiae JAY291]	3.9	No hit

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
901	1-21_TriplEx25	1	gb|EDL10742.1| ERGIC and golgi 2, isoform CRA_b [Mus musculus]	3.9	No hit
902	13-10_TriplEx25	1	gb|AAA62273.1| ORF2 [Trypanosoma brucei]	3.9	No hit
903	5-15_TriplEx25	1	ref|YP_566924.1| phage integrase [Methanococcoides burtonii DSM 6242]	4	No hit
904	34-59_TriplEx25	1	ref|YP_002724418.1| hypothetical protein BBU118A_D17 [Borrelia burgdorferi 118a]	4	No hit
905	34-89_TriplEx25	1	ref|YP_001016594.1| hypothetical protein P9303_05771 [Prochlorococcus marinus]	4	No hit
906	3-32_TriplEx25	1	ref|XP_629561.1| hypothetical protein DDB_G0292668 [Dictyostelium discoideum AX4]	4	No hit
907	11-27_TriplEx25	1	ref|XP_002642886.1| C. briggsae CBR-INFT-1 protein [Caenorhabditis briggsae]	4	No hit
908	3-56_TriplEx25	1	ref|XP_002631557.1| C. briggsae CBR-KQT-3 protein [Caenorhabditis briggsae]	4	No hit
909	34-18_TriplEx25	1	ref|XP_002585980.1| hypothetical protein BRAFLDRAFT_255846 [B. floridae]	4	No hit
910	8-58_TriplEx25	1	ref|XP_002299935.1| predicted protein [Populus trichocarpa]	4	No hit
911	Contig101	6	ref|XP_002262570.1| hypothetical protein [Plasmodium knowlesi]	4	133341l similar to macrophage mennose receptor 1 precursor
912	10-74_TriplEx25	1	ref|XP_002172643.1| VIC ion channel protein cch1 [Schizosaccharomyces japonicus]	4	No hit
913	8-2_TriplEx25	1	ref|XP_001350317.1| hypothetical protein [Plasmodium falciparum 3D7]	4	No hit
914	10-8_TriplEx25	1	ref|XP_001320684.1| ankyrin repeat protein [Trichomonas vaginalis G3]	4	No hit
915	Contig65	2	ref|NP_649625.2| osiris 6 [Drosophila melanogaster]	4	No hit
916	3-50_TriplEx25	1	ref|NP_632352.1| chemotaxis protein [Methanosarcina mazei Go1]	4	No hit
917	10-45_TriplEx25	1	gb|EFA11769.1| hypothetical protein TcasGA2_TC005023 [Tribolium castaneum]	4	No hit
918	34-92_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	4	No hit
919	Contig60	2	emb|CAX64205.1| conserved Plasmodium membrane protein [Plasmodium falciparum]	4	No hit
920	8-20_TriplEx25	1	dbj|BAH58895.1| HMG box transcription factor SoxF(17/18) [Lethenteron japonicum]	4	No hit
921	19-94_TriplEx25	1	ref|ZP_01804293.1| hypothetical protein CdifQ_04001710 [Clostridium difficile QCD-32g58]	4.1	No hit
922	49-10_TriplEx25	1	ref|ZP_01173384.1| peptidyl-tRNA hydrolase [Bacillus sp. NRRL B-14911]	4.1	No hit
923	19-83_TriplEx25	1	ref|YP_082068.1| amino acid ABC transporter [Bacillus cereus]	4.1	No hit
924	49-93_TriplEx25	1	ref|XP_001849894.1| conserved hypothetical protein [Culex quinquefasciatus]	4.1	No hit
925	49-58_TriplEx25	1	ref|XP_001322446.1| PAS domain S-box family protein [Trichomonas vaginalis G3]	4.1	No hit
926	18-11_TriplEx25	1	ref|XP_001259880.1| importin 13, putative [Neosartorya fischeri NRRL 181]	4.1	No hit
927	49-73_TriplEx25	1	ref|XP_001017965.1| putative methyltransferase [Tetrahymena thermophila]	4.1	No hit
928	15-26_TriplEx25	1	ref|ZP_05585810.1| predicted protein [Enterococcus faecalis CH188]	4.3	No hit
929	15-22_TriplEx25	1	ref|ZP_04823480.1| phage protein [Clostridium botulinum E1 str. 'BoNT E Beluga']	4.3	No hit
930	5-34_TriplEx25	1	ref|YP_627178.1| type I restriction enzyme S protein [Helicobacter pylori HPAG1]	4.5	No hit
931	6-45_TriplEx25	1	ref|XP_001950110.1| PREDICTED: similar to ribonuclease iii [Acyrthosiphon pisum]	4.5	No hit
932	42-32_TriplEx25	1	ref|XP_001151770.1| PREDICTED: hypothetical protein [Pan troglodytes]	4.5	No hit
933	6-18_TriplEx25	1	ref|XP_001096454.1| similar to subcommissural organ spondin [Macaca mulatta]	4.5	No hit
934	6-42_TriplEx25	1	ref|NP_504692.2| hypothetical protein K11G9.2 [Caenorhabditis elegans]	4.5	No hit
935	16-42_TriplEx25	1	gb|EET01649.1| Hypothetical protein GL50581_1080 [Giardia intestinalis]	4.5	No hit
936	42-75_TriplEx25	1	gb|EEQ46666.1| hypothetical protein CAWG_05029 [Candida albicans WO-1]	4.5	No hit
937	6-12_TriplEx25	1	gb|ACG52297.1| envelope glycoprotein [Porcine respiratory and reproductive syndrome virus]	4.5	No hit
938	47-14_TriplEx25	1	ref|ZP_05808501.1| Pirin domain protein [Mesorhizobium opportunistum WSM2075]	4.6	No hit
939	47-44_TriplEx25	1	ref|ZP_05313316.1| putative PAS/PAC sensor protein [Geobacter sp. M18]	4.6	No hit
940	47-47_TriplEx25	1	ref|ZP_04806428.1| histidine kinase [Clostridium cellulovorans 743B]	4.6	No hit
941	22-36_TriplEx25	1	ref|ZP_04790445.1| conserved hypothetical protein [Methanocaldococcus infernus ME]	4.6	No hit
942	22-1_TriplEx25	1	ref|ZP_04670145.1| choline/carnitine/betaine transporter family[Clostridiales bacterium]	4.6	No hit
943	40-51_TriplEx25	1	ref|ZP_03929050.1| inner-membrane transport permease ybhR [Acidaminococcus sp. D21]	4.6	No hit
944	32-51_TriplEx25	1	ref|ZP_03458025.1| hypothetical protein BACEGG_00797 [Bacteroides eggerthii]	4.6	No hit
945	38-15_TriplEx25	1	ref|YP_591921.1| short-chain dehydrogenase/reductase SDR [Candidatus Koribacter]	4.6	No hit
946	30-8_TriplEx25	1	ref|YP_196051.1| proline/betaine transporter [Ehrlichia ruminantium str. Gardel]	4.6	No hit
947	35-86_TriplEx25	1	ref|YP_001983232.1| hypothetical protein CJA_2774 [Cellvibrio japonicus Ueda107]	4.6	No hit
948	32-21_TriplEx25	1	ref|YP_001648587.1| NADH dehydrogenase subunit 5 [Ectyoplasia ferox]	4.6	No hit
949	47-22_TriplEx25	1	ref|XP_680245.1| hypothetical protein [Plasmodium berghei strain ANKA]	4.6	No hit
950	32-38_TriplEx25	1	ref|XP_002570557.1| chondroitin sulfate proteoglycan-related [Schistosoma mansoni]	4.6	No hit
951	40-71_TriplEx25	1	ref|XP_002441058.1| hypothetical protein SORBIDRAFT_09g019570 [Sorghum bicolor]	4.6	No hit
952	22-61_TriplEx25	1	ref|XP_002260015.1| hypothetical protein [Plasmodium knowlesi strain H]	4.6	No hit
953	40-66_TriplEx25	1	ref|XP_001991921.1| GH11800 [Drosophila grimshawi]	4.6	No hit
954	35-63_TriplEx25	1	ref|XP_001935839.1| predicted protein [Pyrenophora tritici-repentis Pt-1C-BFP]	4.6	No hit
955	21-33_TriplEx25	1	ref|XP_001930484.1| hypothetical protein PTRG_00151 [Pyrenophora tritici-repentis]	4.6	No hit
956	46-16_TriplEx25	1	ref|XP_001914155.1| hypothetical protein [Entamoeba histolytica HM-1:IMSS]	4.6	No hit
957	47-41_TriplEx25	1	ref|XP_001804123.1| hypothetical protein SNOG_13922 [Phaeosphaeria nodorum SN15]	4.6	No hit
958	37-56_TriplEx25	1	ref|XP_001771070.1| predicted protein [Physcomitrella patens subsp. patens]	4.6	No hit
959	46-40_TriplEx25	1	ref|XP_001733547.1| hypothetical protein [Entamoeba dispar SAW760]	4.6	No hit
960	22-80_TriplEx25	1	ref|XP_001658197.1| serine protease [Aedes aegypti]	4.6	No hit
961	38-34_TriplEx25	1	ref|XP_001606406.1| PREDICTED: similar to synaptic vesicle protein [Nasonia vitripennis]	4.6	No hit
962	40-24_TriplEx25	1	ref|XP_001507112.1| similar to lacrimal androgen-binding protein [Ornithorhynchus anatinus]	4.6	No hit
963	20-92_TriplEx25	1	ref|XP_001443733.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	4.6	No hit
964	45-47_TriplEx25	1	ref|XP_001366979.1| similar to Runt domain containing protein [Monodelphis domestica]	4.6	No hit
965	35-2_TriplEx25	1	ref|XP_001325610.1| hypothetical protein [Trichomonas vaginalis G3]	4.6	No hit
966	22-26_TriplEx25	1	ref|XP_001324406.1| hypothetical protein [Trichomonas vaginalis G3]	4.6	No hit
967	38-48_TriplEx25	1	ref|XP_001324222.1| hypothetical protein [Trichomonas vaginalis G3]	4.6	No hit
968	47-33_TriplEx25	1	ref|NP_241182.1| hypothetical protein BH0316 [Bacillus halodurans C-125]	4.6	No hit
969	32-69_TriplEx25	1	gb|ACP52719.1| eukaryotic initiation factor 2 alpha kinase [Plasmodium berghei]	4.6	No hit
970	47-35_TriplEx25	1	gb|ABV25015.1| beta-galactosidase a-peptide [Cloning vector pTriplEx2]	4.6	No hit
971	35-66_TriplEx25	1	gb|AAB59225.1| NADH dehydrogenase subunit 5 [Trypanosoma brucei]	4.6	No hit
972	30-4_TriplEx25	1	dbj|BAH13086.1| unnamed protein product [Homo sapiens]	4.6	No hit
973	40-58_TriplEx25	1	dbj|BAA83309.1| maturase [Dracaena angustifolia]	4.6	No hit
974	24-57_TriplEx25	1	ref|ZP_02948398.1| putative metallo-beta-lactamase family protein [Clostridium butyricum]	4.7	No hit
975	24-75_TriplEx25	1	ref|XP_002610740.1| hypothetical protein BRAFLDRAFT_90927 [B. floridae]	4.7	No hit
976	25-26_TriplEx25	1	ref|NP_647965.2| CG4835 [Drosophila melanogaster]	4.7	No hit
977	27-2_TriplEx25	1	ref|ZP_02861304.1| hypothetical protein ANASTE_00504 [Anaerofustis stercorihominis]	4.9	133228l glucokinase regulator
978	28-69_TriplEx25	1	ref|ZP_01666400.1| major facilitator superfamily MFS_1 [Thermosinus carboxydivorans]	4.9	No hit
979	26-52_TriplEx25	1	ref|YP_302721.1| hypothetical protein Ecaj_0072 [Ehrlichia canis str. Jake]	4.9	No hit
980	26-8_TriplEx25	1	ref|YP_001673041.1| formate dehydrogenase, alpha subunit [Shewanella halifaxensis]	4.9	No hit
981	27-9_TriplEx25	1	ref|XP_001378932.1| PREDICTED: hypothetical protein [Monodelphis domestica]	4.9	No hit
982	28-25_TriplEx25	1	gb|EER05546.1| conserved hypothetical protein [Perkinsus marinus ATCC 50983]	4.9	No hit
983	26-27_TriplEx25	1	emb|CAO98872.1| hypothetical protein [Nakaseomyces delphensis]	4.9	No hit
984	2-27_TriplEx25	1	ref|ZP_03487854.1| hypothetical protein EUBIFOR_00419 [Eubacterium biforme DSM 3989]	5.1	No hit
985	2-50_TriplEx25	1	ref|ZP_01906283.1| NCS1 nucleoside transporter family protein [Plesiocystis pacifica]	5.1	No hit
986	17-77_TriplEx25	1	ref|ZP_01254194.1| gamma-glutamyl carboxylase-like protein [Psychroflexus torquis]	5.1	No hit
987	1-77_TriplEx25	1	ref|ZP_01254194.1| gamma-glutamyl carboxylase-like protein [Psychroflexus torquis]	5.1	No hit
988	1-89_TriplEx25	1	ref|YP_417839.1| membrane-spanning protein [Staphylococcus aureus RF122]	5.1	No hit
989	2-95_TriplEx25	1	ref|XP_790542.2| PREDICTED: similar to Grp94 neighboring nucleotidase [S. purpuratus]	5.1	No hit
990	12-88_TriplEx25	1	ref|XP_002569255.1| Pc21g22880 [Penicillium chrysogenum Wisconsin 54-1255]	5.1	No hit
991	2-28_TriplEx25	1	ref|XP_001309051.1| surface antigen BspA-like [Trichomonas vaginalis G3]	5.1	No hit
992	9-52_TriplEx25	1	ref|ZP_05733600.2| putative outer membrane ferric siderophore [Dialister invisus]	5.2	No hit
993	7-33_TriplEx25	1	ref|ZP_01859368.1| Glutathione peroxidase [Bacillus sp. SG-1]	5.2	No hit
994	33-79_TriplEx25	1	ref|ZP_01447312.1| 3-deoxy-D-manno-octulosonic-acid transferase [alpha proteobacterium]	5.2	No hit
995	3-48_TriplEx25	1	ref|YP_002782702.1| hypothetical protein ROP_55100 [Rhodococcus opacus B4]	5.2	No hit
996	3-30_TriplEx25	1	ref|YP_002323227.1| signal transduction histidine kinase [Bifidobacterium longum]	5.2	No hit
997	34-91_TriplEx25	1	ref|YP_001884677.1| mannosyltransferase B [Clostridium botulinum B str. Eklund 17B]	5.2	No hit
998	34-66_TriplEx25	1	ref|YP_001256633.1| hypothetical protein MAG_4950 [Mycoplasma agalactiae PG2]	5.2	No hit
999	11-9_TriplEx25	1	ref|XP_392454.2| PREDICTED: similar to alpha-2-macroglobulin-like 1 [Apis mellifera]	5.2	No hit
1000	34-57_TriplEx25	1	ref|XP_002366154.1| hypothetical protein TGME49_024800 [Toxoplasma gondii ME49]	5.2	No hit

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
1001	4-45_TriplEx25	1	ref|XP_001987400.1| GH21902 [Drosophila grimshawi]	5.2	No hit
1002	34-48_TriplEx25	1	ref|XP_001737768.1| PH domain leucine-rich repeat protein phosphatase [Entamoeba dispar]	5.2	No hit
1003	9-29_TriplEx25	1	ref|XP_001448245.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	5.2	No hit
1004	3-4_TriplEx25	1	ref|XP_001323341.1| hypothetical protein [Trichomonas vaginalis G3]	5.2	No hit
1005	7-6_TriplEx25	1	ref|XP_001144824.1| PREDICTED: pepsinogen 5 [Pan troglodytes]	5.2	No hit
1006	11-72_TriplEx25	1	gb|EFA04887.1| hypothetical protein TcasGA2_TC014949 [Tribolium castaneum]	5.2	No hit
1007	10-94_TriplEx25	1	gb|EEU41912.1| predicted protein [Nectria haematococca mpVI 77-13-4]	5.2	No hit
1008	10-59_TriplEx25	1	emb|CAN61210.1| hypothetical protein [Vitis vinifera]	5.2	No hit
1009	49-13_TriplEx25	1	ref|ZP_05840845.1| Radical SAM domain protein [Ferroglobus placidus DSM 10642]	5.4	No hit
1010	19-13_TriplEx25	1	ref|ZP_04568673.1| predicted protein [Fusobacterium mortiferum ATCC 9817]	5.4	No hit
1011	49-14_TriplEx25	1	ref|YP_435621.1| HD-GYP domain-containing protein [Hahella chejuensis KCTC 2396]	5.4	No hit
1012	48-25_TriplEx25	1	ref|YP_393853.1| pseudouridylate synthase [Sulfurimonas denitrificans DSM 1251]	5.4	No hit
1013	19-56_TriplEx25	1	ref|YP_002382909.1| disulfide bond formation protein B [Escherichia fergusonii]	5.4	No hit
1014	18-57_TriplEx25	1	ref|YP_001942216.1| conserved hypothetical protein [Chlorobium limicola DSM 245]	5.4	No hit
1015	17-32_TriplEx25	1	ref|YP_001906957.1| stationary phase inducible protein CsiE [Erwinia tasmaniensis]	5.4	No hit
1016	48-96_TriplEx25	1	ref|XP_641828.1| transcription initiation factor IIF subunit alpha [Dictyostelium discoideum]	5.4	No hit
1017	49-9_TriplEx25	1	ref|XP_002600673.1| hypothetical protein BRAFLDRAFT_67734 [B. floridae]	5.4	No hit
1018	50-57_TriplEx25	1	ref|XP_001441956.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	5.4	No hit
1019	49-7_TriplEx25	1	ref|NP_149817.1| 354L [Invertebrate iridescent virus 6]	5.4	No hit
1020	19-75_TriplEx25	1	gb|EEC80074.1| hypothetical protein OsI_21793 [Oryza sativa Indica Group]	5.4	No hit
1021	17-53_TriplEx25	1	gb|ACQ72595.1| NADH dehydrogenase subunit 5 [Anopheles paltrinierii]	5.4	No hit
1022	49-35_TriplEx25	1	gb|ACN27574.1| unknown [Zea mays]	5.4	No hit
1023	44-24_TriplEx25	1	ref|ZP_04778602.1| group 1 glycosyl transferase [Sphingobacterium spiritivorum]	5.9	No hit
1024	5-50_TriplEx25	1	ref|ZP_04620736.1| hypothetical protein yaldo0001_34050 [Yersinia aldovae ATCC 35236]	5.9	No hit
1025	42-61_TriplEx25	1	ref|ZP_03972856.1| cytochrome C oxidase polypeptide I [Corynebacterium glucuronolyticum]	5.9	No hit
1026	6-89_TriplEx25	1	ref|ZP_01859493.1| hypothetical protein BSG1_11146 [Bacillus sp. SG-1]	5.9	No hit
1027	42-69_TriplEx25	1	ref|YP_858385.1| ABC-type sugar transport system component [Aeromonas hydrophila]	5.9	No hit
1028	42-73_TriplEx25	1	ref|YP_001514734.1| hypothetical protein AM1_0361 [Acaryochloris marina MBIC11017]	5.9	No hit
1029	6-64_TriplEx25	1	ref|YP_001260503.1| ATP synthase F0 subunit 8 [Ammotragus lervia]	5.9	No hit
1030	42-89_TriplEx25	1	ref|XP_001899220.1| hypothetical protein [Brugia malayi]	5.9	No hit
1031	5-32_TriplEx25	1	ref|XP_001691705.1| ligand-gated ion channel [Chlamydomonas reinhardtii]	5.9	No hit
1032	42-14_TriplEx25	1	ref|XP_001470486.1| hypothetical protein [Leishmania infantum JPCM5]	5.9	No hit
1033	6-77_TriplEx25	1	ref|XP_001443362.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	5.9	No hit
1034	41-25_TriplEx25	1	gb|EEU35466.1| hypothetical protein NECHADRAFT_101975 [Nectria haematococca]	5.9	No hit
1035	43-41_TriplEx25	1	gb|EEH50340.1| C-x8-C-x5-C-x3-H type zinc finger protein [Paracoccidioides brasiliensis]	5.9	No hit
1036	42-91_TriplEx25	1	gb|EEH50340.1| C-x8-C-x5-C-x3-H type zinc finger protein [Paracoccidioides brasiliensis]	5.9	No hit
1037	38-5_TriplEx25	1	ref|XP_686943.3| PREDICTED: similar to restin, partial [Danio rerio]	6	No hit
1038	25-16_TriplEx25	1	ref|ZP_04874239.1| Integral membrane protein [Aciduliprofundum boonei T469]	6.1	No hit
1039	29-88_TriplEx25	1	ref|YP_001909904.1| hypothetical protein HPSH_02120 [Helicobacter pylori Shi470]	6.1	No hit
1040	30-2_TriplEx25	1	ref|YP_001863733.1| hypothetical protein Bphy_7807 [Burkholderia phymatum STM815]	6.1	No hit
1041	23-35_TriplEx25	1	ref|XP_454901.1| unnamed protein product [Kluyveromyces lactis]	6.1	No hit
1042	37-71_TriplEx25	1	ref|XP_001718531.2| PREDICTED: hypothetical protein [Homo sapiens]	6.1	No hit
1043	25-50_TriplEx25	1	ref|XP_001682696.1| hypothetical protein [Leishmania major strain Friedlin]	6.1	No hit
1044	31-68_TriplEx25	1	ref|XP_001520135.1| similar to V1R pheromone receptor-like protein [Ornithorhynchus anatinus]	6.1	No hit
1045	35-33_TriplEx25	1	ref|XP_001418256.1| predicted protein [Ostreococcus lucimarinus CCE9901]	6.1	No hit
1046	29-16_TriplEx25	1	ref|XP_001211937.1| conserved hypothetical protein [Aspergillus terreus NIH2624]	6.1	No hit
1047	Contig164	2	ref|XP_002586282.1| hypothetical protein BRAFLDRAFT_254260 [B. floridae]	6.3	No hit
1048	26-17_TriplEx25	1	ref|XP_635134.1| hypothetical protein DDB_G0291308 [Dictyostelium discoideum AX4]	6.4	No hit
1049	27-47_TriplEx25	1	ref|XP_001963738.1| GF21101 [Drosophila ananassae]	6.4	No hit
1050	27-17_TriplEx25	1	gb|EEE34125.1| conserved hypothetical protein [Toxoplasma gondii VEG]	6.4	No hit
1051	26-86_TriplEx25	1	gb|AAU83891.1| hypothetical protein GZ34H9_9 [uncultured archaeon GZfos34H9]	6.4	No hit
1052	27-49_TriplEx25	1	emb|CAL57412.1| CMV 1a interacting protein 1 (ISS) [Ostreococcus tauri]	6.4	No hit
1053	27-77_TriplEx25	1	emb|CAG04938.1| unnamed protein product [Tetraodon nigroviridis]	6.4	112416l similar von Willebrand factor type A, EGF and pentraxin domain
1054	26-2_TriplEx25	1	dbj|BAC00935.1| S2-RNase [Solanum chilense]	6.4	No hit
1055	Contig71	2	ref|ZP_05920545.1| conserved hypothetical protein [Pasteurella dagmatis ATCC 43325]	6.6	No hit
1056	15-69_TriplEx25	1	ref|ZP_02862576.1| hypothetical protein ANASTE_01795 [Anaerofustis stercorihominis]	6.6	No hit
1057	12-96_TriplEx25	1	ref|XP_732070.1| hypothetical protein [Plasmodium chabaudi chabaudi]	6.6	No hit
1058	Contig100	2	ref|XP_002577427.1| nyd-sp30 tubulin tyrosine ligase-related [Schistosoma mansoni]	6.6	No hit
1059	12-12_TriplEx25	1	ref|XP_001918315.1| PREDICTED: similar to NLRP3 protein [Equus caballus]	6.6	No hit
1060	16-53_TriplEx25	1	ref|XP_001896898.1| TATA binding protein associated factor [Brugia malayi]	6.6	No hit
1061	Contig85	2	ref|NP_001072316.1| glypican 1 [Xenopus (Silurana) tropicalis]	6.6	No hit
1062	16-3_TriplEx25	1	gb|ACO09627.1| Pannexin-1 [Osmerus mordax]	6.6	No hit
1063	Contig98	2	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	6.6	No hit
1064	1-2_TriplEx25	1	ref|ZP_03946389.1| conserved hypothetical protein [Atopobium vaginae DSM 15829]	6.7	No hit
1065	Contig111	2	ref|NP_049597.1| orf1386 [Tetrahymena pyriformis]	6.7	No hit
1066	2-26_TriplEx25	1	gb|EAZ32406.1| hypothetical protein OsJ_16617 [Oryza sativa Japonica Group]	6.7	No hit
1067	33-88_TriplEx25	1	ref|ZP_05916052.1| hypothetical protein HMPREF6745_0005 [Prevotella sp.]	6.8	No hit
1068	4-62_TriplEx25	1	ref|ZP_05084158.1| chorismate binding enzyme [Pseudovibrio sp. JE062]	6.8	No hit
1069	4-24_TriplEx25	1	ref|ZP_03611258.1| type VI secretion protein, family [Campylobacter rectus RM3267]	6.8	No hit
1070	11-26_TriplEx25	1	ref|ZP_03489861.1| hypothetical protein EUBIFOR_02457 [Eubacterium biforme DSM 3989]	6.8	No hit
1071	4-14_TriplEx25	1	ref|YP_002725713.1| NADH dehydrogenase subunit 5 [Ancylostoma caninum]	6.8	No hit
1072	34-60_TriplEx25	1	ref|YP_002607878.1| folylpolyglutamate synthetase [Nautilia profundicola AmH]	6.8	No hit
1073	8-45_TriplEx25	1	ref|XP_714267.1| hypothetical protein CaO19.4096 [Candida albicans SC5314]	6.8	No hit
1074	8-83_TriplEx25	1	ref|XP_586303.1| PREDICTED: similar to C-type lectin domain family [Bos taurus]	6.8	No hit
1075	3-44_TriplEx25	1	ref|XP_002261514.1| hypothetical protein [Plasmodium knowlesi strain H]	6.8	No hit
1076	4-37_TriplEx25	1	ref|XP_001743116.1| hypothetical protein [Monosiga brevicollis MX1]	6.8	No hit
1077	11-23_TriplEx25	1	ref|XP_001527334.1| predicted protein [Lodderomyces elongisporus NRRL YB-4239]	6.8	No hit
1078	34-33_TriplEx25	1	ref|XP_001493705.2| PREDICTED: similar to odorant receptor [Equus caballus]	6.8	No hit
1079	9-81_TriplEx25	1	ref|XP_001470858.1| voltage and ligand gated potassium channel [Tetrahymena thermophila]	6.8	No hit
1080	33-45_TriplEx25	1	ref|XP_001329659.1| PIKK family atypical protein kinase [Trichomonas vaginalis G3]	6.8	No hit
1081	34-25_TriplEx25	1	ref|XP_001317831.1| ankyrin repeat protein [Trichomonas vaginalis G3]	6.8	No hit
1082	7-66_TriplEx25	1	ref|XP_001008079.1| hypothetical protein TTHERM_00005950 [Tetrahymena thermophila]	6.8	No hit
1083	Contig41	4	gb|ACX48286.1| cytochrome oxidase subunit II [Pinnaspis uniloba]	6.8	No hit
1084	4-72_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	6.8	No hit
1085	7-69_TriplEx25	1	emb|CAD67931.1| putative A-ATPase C-subunit [Thermotoga sp. RQ2]	6.8	No hit
1086	17-12_TriplEx25	1	ref|ZP_02156672.1| tRNA (guanine-N(7)-)-methyltransferase [Shewanella benthica KT99]	7	No hit
1087	48-7_TriplEx25	1	ref|XP_627833.1| hypothetical protein [Cryptosporidium parvum Iowa II]	7	No hit
1088	49-80_TriplEx25	1	ref|XP_002639024.1| C. briggsae CBR-RPL-24.1 protein [Caenorhabditis briggsae]	7	No hit
1089	19-61_TriplEx25	1	ref|XP_002603982.1| hypothetical protein BRAFLDRAFT_71730 [B. floridae]	7	No hit
1090	49-74_TriplEx25	1	ref|XP_002532673.1| Xyloglucan endotransglucosylase [Ricinus communis]	7	No hit
1091	18-78_TriplEx25	1	ref|XP_001593006.1| hypothetical protein SS1G_05928 [Sclerotinia sclerotiorum 1980]	7	No hit
1092	17-42_TriplEx25	1	ref|NP_760362.2| Actin-like ATPase involved in cell morphogenesis [Vibrio vulnificus]	7	No hit
1093	17-20_TriplEx25	1	ref|NP_348620.1| indolepyruvate ferredoxin oxidoreductase [Clostridium acetobutylicum]	7	No hit
1094	48-28_TriplEx25	1	ref|NP_064793.1| hypothetical protein AMV011 [Amsacta moorei entomopoxvirus 'L']	7	No hit
1095	48-50_TriplEx25	1	gb|EER42505.1| fungal specific transcription factor [Ajellomyces capsulatus]	7	No hit
1096	19-36_TriplEx25	1	gb|ACZ56432.1| fusion protein [Canine distemper virus]	7	No hit
1097	48-13_TriplEx25	1	dbj|BAE47410.1| delta antigen [Hepatitis delta virus]	7	No hit
1098	15-8_TriplEx25	1	ref|XP_001432879.1| hypothetical protein [Paramecium tetraurelia strain d4-2]	7.5	No hit
1099	41-31_TriplEx25	1	ref|ZP_01883232.1| hypothetical protein PBAL39_09381 [Pedobacter sp. BAL39]	7.7	No hit
1100	41-34_TriplEx25	1	ref|YP_001878235.1| conserved hypothetical protein, membrane [Akkermansia muciniphila]	7.7	No hit

Appendix. Continued.

	Clone number	Numbe of clone	NCBI nr-BlastX	E-value	Bf genome-BlastX
1101	6-29_TriplEx25	1	ref|YP_001212679.1| hypothetical protein PTH_2129 [Pelotomaculum thermopropionicum]	7.7	No hit
1102	42-21_TriplEx25	1	ref|XP_745899.1| hypothetical protein [Plasmodium chabaudi chabaudi]	7.7	No hit
1103	6-56_TriplEx25	1	ref|XP_002291488.1| predicted protein [Thalassiosira pseudonana CCMP1335]	7.7	No hit
1104	41-67_TriplEx25	1	gb|ACU78325.1| conserved hypothetical protein [Mycoplasma mycoides]	7.7	No hit
1105	43-83_TriplEx25	1	db|BAF91637.1| M polyprotein [Tinaroo virus]	7.7	No hit
1106	20-3_TriplEx25	1	sp|P53880.1|Putative uncharacterized protein [Saccharomyces cerevisiae]	7.9	No hit
1107	36-46_TriplEx25	1	ref|ZP_05779130.1| ATP-dependent DNA helicase RecG [Dialister invisus DSM 15470]	7.9	No hit
1108	24-73_TriplEx25	1	ref|ZP_04807445.1| Methyltransferase type 11 [Clostridium cellulovorans 743B]	7.9	No hit
1109	20-18_TriplEx25	1	ref|ZP_01874465.1| probable ECF sigma factor [Lentisphaera araneosa HTCC2155]	7.9	No hit
1110	24-13_TriplEx25	1	ref|ZP_01101263.1| GTP pyrophosphokinase [Congregibacter litoralis KT71]	7.9	No hit
1111	30-28_TriplEx25	1	ref|YP_942739.1| sulphate transporter [Psychromonas ingrahamii 37]	7.9	No hit
1112	20-79_TriplEx25	1	ref|YP_002444317.1| sensor histidine kinase [Bacillus cereus G9842]	7.9	No hit
1113	24-76_TriplEx25	1	ref|XP_765161.1| hypothetical protein [Theileria parva strain Muguga]	7.9	No hit
1114	24-51_TriplEx25	1	ref|XP_729524.1| hypothetical protein [Plasmodium yoelii yoelii str. 17XNL]	7.9	No hit
1115	25-54_TriplEx25	1	ref|XP_002482852.1| conserved hypothetical protein [Talaromyces stipitatus ATCC 10500]	7.9	No hit
1116	25-14_TriplEx25	1	ref|XP_002156246.1| PREDICTED: similar to Protein RSM22 homolog [Hydra magnipapillata]	7.9	No hit
1117	21-64_TriplEx25	1	ref|XP_001013225.1| hypothetical protein TTHERM_00295930 [Tetrahymena thermophila]	7.9	No hit
1118	22-55_TriplEx25	1	gb|ACQ58565.1| Myeloid cell surface antigen CD33 precursor [Anoplopoma fimbria]	7.9	No hit
1119	20-6_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	7.9	No hit
1120	30-41_TriplEx25	1	gb|ABS90170.2| hypothetical protein A1S_3745 [Acinetobacter baumannii ATCC 17978]	7.9	No hit
1121	22-71_TriplEx25	1	emb|CAB76451.2| acetyl xylan esterase [Bacillus pumilus]	7.9	No hit
1122	20-57_TriplEx25	1	emb|CAA67813.1| extensin-like protein Dif10 [Solanum lycopersicum]	7.9	No hit
1123	25-18_TriplEx25	1	ref|XP_001717734.2| PREDICTED: hypothetical protein [Homo sapiens]	8.2	No hit
1124	28-28_TriplEx25	1	ref|ZP_05759429.1| hypothetical protein BacD2_14193 [Bacteroides sp. D2]	8.4	No hit
1125	27-25_TriplEx25	1	ref|YP_271259.1| aminotransferase [Colwellia psychrerythraea 34H]	8.4	No hit
1126	26-55_TriplEx25	1	ref|XP_002592696.1| hypothetical protein BRAFLDRAFT_67135 [B. floridae]	8.4	No hit
1127	26-48_TriplEx25	1	ref|XP_001712572.1| taf90 [Hemiselmis andersenii]	8.4	No hit
1128	13-75_TriplEx25	1	ref|XP_675894.1| hypothetical protein [Plasmodium berghei strain ANKA]	8.6	No hit
1129	16-32_TriplEx25	1	ref|XP_001327175.1| surface antigen BspA-like [Trichomonas vaginalis G3]	8.6	No hit
1130	12-78_TriplEx25	1	gb|EDL77485.1| acid phosphatase-like 2 [Rattus norvegicus]	8.6	No hit
1131	16-28_TriplEx25	1	gb|ACB70387.1| hypothetical protein [Ornithodoros coriaceus]	8.6	No hit
1132	Contig77	3	ref|ZP_05081190.1| sun protein [beta proteobacterium KB13]	8.7	No hit
1133	2-38_TriplEx25	1	ref|XP_676200.1| hypothetical protein [Plasmodium berghei strain ANKA]	8.7	No hit
1134	1-6_TriplEx25	1	ref|XP_002022899.1| GL16474 [Drosophila persimilis]	8.7	No hit
1135	2-76_TriplEx25	1	ref|XP_001614551.1| hypothetical protein [Plasmodium vivax Sal-1]	8.7	No hit
1136	2-5_TriplEx25	1	ref|NP_976173.1| NADH dehydrogenase subunit 4 [Biomphalaria glabrata]	8.7	No hit
1137	34-63_TriplEx25	1	ref|YP_001238862.1| putative ferredoxin--NAD(+) reductase [Bradyrhizobium sp. BTAi1]	8.8	No hit
1138	34-46_TriplEx25	1	ref|XP_002199386.1| PREDICTED: WD repeat domain 67 [Taeniopygia guttata]	8.8	No hit
1139	33-14_TriplEx25	1	ref|XP_002125314.1| similar to NFX1-type zinc finger-containing protein [C. intestinalis]	8.8	No hit
1140	11-48_TriplEx25	1	ref|ZP_03008330.1| hypothetical protein BACCOP_00169 [Bacteroides coprocola]	8.9	No hit
1141	11-75_TriplEx25	1	ref|ZP_01869192.1| ribonucleotide-diphosphate reductase alpha subunit [Vibrio shilonii]	8.9	No hit
1142	3-51_TriplEx25	1	ref|ZP_01215591.1| putative regulatory protein [Psychromonas sp. CNPT3]	8.9	No hit
1143	4-73_TriplEx25	1	ref|XP_002446605.1| hypothetical protein SORBIDRAFT_06g018790 [Sorghum bicolor]	8.9	No hit
1144	8-56_TriplEx25	1	ref|XP_001869523.1| conserved hypothetical protein [Culex quinquefasciatus]	8.9	No hit
1145	8-63_TriplEx25	1	ref|XP_001341690.2| PREDICTED: similar to 6720455I24Rik homolog [Danio rerio]	8.9	No hit
1146	7-26_TriplEx25	1	gb|ABW38935.1| putative CAD trifunctional protein [Cydia pomonella]	8.9	No hit
1147	7-52_TriplEx25	1	gb|AAM28670.1| NADH dehydrogenase subunit F [Scabiosa columbaria]	8.9	No hit
1148	11-17_TriplEx25	1	emb|CAL54426.1| putative callose synthase 1 catalytic subunit (ISS) [Ostreococcus tauri]	8.9	No hit
1149	18-95_TriplEx25	1	ref|ZP_03166450.1| hypothetical protein RUMLAC_00096 [Ruminococcus lactaris]	9.2	No hit
1150	49-53_TriplEx25	1	ref|XP_737331.1| hypothetical protein [Plasmodium chabaudi chabaudi]	9.2	No hit
1151	48-8_TriplEx25	1	ref|XP_558523.3| AGAP005135-PA [Anopheles gambiae str. PEST]	9.2	No hit
1152	18-68_TriplEx25	1	ref|XP_002609538.1| hypothetical protein BRAFLDRAFT_241410 [B. floridae]	9.2	No hit
1153	49-55_TriplEx25	1	ref|XP_002596057.1| hypothetical protein BRAFLDRAFT_203157 [B. floridae]	9.2	No hit
1154	49-67_TriplEx25	1	ref|XP_002124265.1| PREDICTED: similar to plexin-B1/SEP receptor, partial [C. intestinalis]	9.2	No hit
1155	48-39_TriplEx25	1	ref|XP_002028154.1| GL15388 [Drosophila persimilis]	9.2	No hit
1156	49-34_TriplEx25	1	ref|XP_001716895.2| PREDICTED: chromosome 10 open reading frame 112 [Homo sapiens]	9.2	No hit
1157	18-42_TriplEx25	1	ref|NP_982740.1| ABL207Wp [Ashbya gossypii ATCC 10895]	9.2	No hit
1158	17-79_TriplEx25	1	ref|NP_499332.2| hypothetical protein Y45F3A.8 [Caenorhabditis elegans]	9.2	No hit
1159	48-31_TriplEx25	1	ref|NP_001009454.1| growth hormone releasing hormone receptor [Ovis aries]	9.2	No hit
1160	19-31_TriplEx25	1	gb|EEU44533.1| hypothetical protein NECHADRAFT_96455 [Nectria haematococca]	9.2	No hit
1161	19-96_TriplEx25	1	gb|AAM65428.1| putative C-4 sterol methyl oxidase [Arabidopsis thaliana]	9.2	No hit
1162	48-16_TriplEx25	1	emb|CAM77595.1| hypothetical protein [Magnetospirillum gryphiswaldense MSR-1]	9.2	No hit
1163	49-59_TriplEx25	1	emb|CAG24999.2| conserved Plasmodium membrane protein [Plasmodium chabaudi]	9.2	No hit
1164	14-66_TriplEx25	1	emb|CAN60854.1| hypothetical protein [Vitis vinifera]	9.7	No hit